# INTRODUCTION
# TO
# ECONOMICS

| | |
|---|---|
| SYDNEY<br>MELBOURNE | BUTTERWORTH & CO (AUSTRALIA) LTD |
| BOMBAY | BUTTERWORTH & CO (INDIA) LTD |
| TORONTO | BUTTERWORTH & CO (CANADA) LTD |
| WELLINGTON (N Z)<br>AUCKLAND (N Z) | BUTTERWORTH & CO (AUSTRALIA) LTD |
| DURBAN | BUTTERWORTH & CO (AFRICA) LTD |

# INTRODUCTION

# TO

# ECONOMICS

BY

ALEC CAIRNCROSS

*Ph.D. (Cantab)*

Lecturer in Political Economy in the
University of Glasgow

LONDON

BUTTERWORTH & CO. (PUBLISHERS) LTD.

BELL YARD, TEMPLE BAR

1944

To

The Cambridge Economists

of all countries.

# PREFACE

With the exception of a few paragraphs, chiefly in Chapters 20, 28 and 29, a first draft of this book was complete when war broke out. The book was then laid aside for three years and when eventually it came to be prepared for publication, the drastic revision which it would normally have undergone was out of the question. Amongst other things it was impossible to remedy an obvious lack of balance between the treatment of problems of industrial organisation and the treatment of problems of unemployment and international trade.

In form and matter the book follows conventional lines and makes no claim to originality. It is written for those who, anxious to understand the complexities of the economic order and the causes of economic disorder, find it difficult to master the economist's technique of thinking and hesitate to plunge too abruptly into analytical deeps. The early chapters move at a fairly slow tempo so as to allow the reader to accustom himself to the atmosphere of the subject; but there is no attempt to over-simplify, and no deliberate evasion of fundamental issues.

My major obligation is expressed in the dedication: I hope that I have been not unfaithful to the spirit of those who taught and studied in Cambridge ten years ago. I also owe a great debt to Dr. A. L. Macfie for his constant encouragement and suggestions; and to Messrs. Neilson and Palmer of the School of Accountancy for the instigation and castigation that made me (with some hesitation) start, and (with more hesitation) finish the work. Amongst others who have been kind enough to offer criticisms of parts of the manuscript I would like to thank Dr. H. W. Singer, Dr. A. S. Cairncross, Mr. A. D. K. Owen and Mr. P. T. Bauer.

This book is, of course, in no way associated with my official duties in the Ministry of Aircraft Production.

ALEC CAIRNCROSS.

London, May 1944.

# CONTENTS


| Chapter | | Page |
|---|---|---|
| | PART I: INTRODUCTORY | |
| 1 | The Economic Problem | 1 |
| 2 | Some Preliminary Conceptions | 13 |
| 3 | The Existing Resources: The Factors or Agents of Production | 24 |
| | PART II: THE STRUCTURE OF INDUSTRY | |
| 4 | Specialisation: (1) Division of Labour | 43 |
| 5 | Specialisation: (2) Localisation of Industry | 52 |
| 6 | Large-Scale Production | 61 |
| 7 | Small-Scale Production | 80 |
| 8 | The Finance of Large-Scale Production | 89 |
| 9 | The Growth of Large-Scale Production | 106 |
| | PART III: SUPPLY AND DEMAND | |
| 10 | The Existing Order: a Preliminary Survey | 138 |
| 11 | Pricing | 143 |
| 12 | Supply and Demand | 151 |
| 13 | Demand | 160 |
| 14 | Supply: (1) Competition and Monopoly | 171 |
| 15 | Supply: (2) Cost in Relation to Output | 188 |
| 16 | Supply: (3) Cost in Relation to Time | 196 |
| 17 | Inter-relationships of Supply and Demand | 202 |
| 18 | Social Aspects of Pricing | 209 |
| | PART IV: THE DISTRIBUTION OF INCOME | |
| 19 | The Distribution of Income | 216 |
| 20 | Wages | 236 |
| 21 | Interest and Profit | 251 |
| | PART V.: INTERNATIONAL TRADE | |
| 22 | The Theory of International Trade | 279 |
| | PART VI: MONEY AND BANKING | |
| 23 | Money | 292 |
| 24 | Banking | 302 |
| 25 | The Supply of Money | 325 |
| 26 | The Value of Money | 333 |
| 27 | Money and Foreign Exchange | 357 |
| 28 | Unemployment and the Trade Cycle | 374 |
| | PART VII: THE STATE AND THE ECONOMIC ORDER | |
| 29 | The State and Industry | 395 |
| | INDEX | |

# PART I—INTRODUCTORY

## CHAPTER 1

## THE ECONOMIC PROBLEM

ECONOMICS is first and foremost a study of human behaviour. So, too, are many other subjects—ethics, politics, and psychology, for example. What aspect of human behaviour is the special point of departure of economics?

A preliminary answer is that economics studies the part played by *money* in human affairs. It is about money and the things that can be bought and sold for money, from bales of cotton to a cup of tea. It is about the getting and spending of money: how men earn a living and what sort of living it is that they earn. And it is about the way in which money affects our way of life and our outlook on life.

If we probe a little deeper, however, we find that economics is really not so much about money as about some things which are implied in the use of money. Three of these—exchange, scarcity, and choice—are of special importance. Let us take them in turn.

### 1. EXCHANGE.

Money implies exchange. It is in fact the medium of exchange. In a primitive community, where exchanges are rare, we can dispense with money and resort to direct barter. Money is unnecessary so long as we are at the stage of trying to satisfy all our wants by our own efforts, growing our own wheat, milling our own flour, baking our own bread, and only now and again exchanging, say, wheat for a ploughshare or a calf for a millstone. But immediately we begin to specialise, and cease to produce goods for our own use, money becomes indispensable if exchanges are to take place smoothly. Exchange becomes triangular—we convert goods into money and money into other goods, instead of simply bartering goods for goods. If exchanges did not take place in this apparently circuitous way, no one who specialised in making bricks or bowler hats would relish a morning's shopping. The grocer might have no use for bricks; and match-sellers would hesitate to accept the hundredth part of a bowler hat. A walletful of money goes so much further than other walletfuls!

Nowadays, therefore, exchange rarely takes the form of direct barter. Instead, we do business with money. We buy what we want with money, sell for money, fix prices in terms of money, are paid our wages, salaries, or dividends in money, save money, and measure our wealth in money. But the problems which present themselves to us in terms of money are exactly similar to the problems raised by direct barter. There is a surface difference between money-exchange and barter-exchange, but no difference in principle. Economics, there-

fore, does not limit itself only to money-problems but studies exchange-problems of all kinds. It is, in fact, about exchange rather than about money, for exchange underlies the use of money.

**Exchange implies Interdependence.**—When we exchange, we have stopped being self-sufficing and have become dependent on those from whom we buy and to whom we sell. Our fortunes are linked with theirs. If they are poor or unemployed then we are likely to be in danger of poverty and unemployment ourselves. Famine and flood in one part of the world can create scarcity and distress thousands of miles away by cutting off supplies of foodstuffs and raw materials. We are all within the circle of exchange. Yet this interdependence rarely occurs to us: it is so easy to overlook the implications of exchange. Consider, for example, some everyday event like the purchase of a packet of cigarettes. I take from my pocket a small piece of metal—probably Mexican silver alloyed with Canadian nickel—and offer it to a total stranger who accepts it with alacrity. In exchange, I receive a cardboard packet whose contents are the product of workers from all over the globe—Norwegian lumbermen, Turkish peasants, Malayan tin-miners, American inventors. I draw also on the services of British workers scattered over the country. The packing of the cigarettes has been done in Bristol; the cellophane wrapper and silver paper come from London; the paper round the cigarettes from Swindon; the stiffener, or cigarette card, from Glasgow. But about all these workers, through whose efforts I am able to smoke my cigarettes, I am amazingly ignorant. I do not trouble to inquire whether they include cannibals, racketeers, and Jew-baiters; whether they are mean, grasping, or dissolute; or whether their daily earnings are less than 1d. or over £100. Their creed, their way of living, their income, the colour of their skin, do not interest me. I can drive my bargain with them without even knowing that they exist. The cash-nexus that binds us is the loosest of bonds. It leaves me free to pursue my own interest, undeterred by any sense of moral obligation to other workers as fellow-citizens. They satisfy my wants and earn the means of satisfying theirs. And that, to most of us, might seem to be the end of the matter.

But *not* to the economist. It is precisely these exchange-bargains which he sets out to investigate. Why, he asks, do people exchange at all? What advantage does society reap from leaving people free to satisfy their wants by exchange? When is exchange fair and when unfair? Is it in the social interest that exchanges dictated by mutual self-interest should be left unregulated by the State? Or, if regulation is desirable, on what principles should the State intervene?

## 2. Scarcity.

The use of money implies scarcity. Money itself must be scarce or it will cease to be used. If the supply of money is increased without limit it will soon lose value and in the end no one will accept it. Whatever passes as money, therefore, must necessarily be scarce. So also—

and this is the important point—must be the things that money will buy. We only exchange one scarce thing for another. We do not pay for air and earth and water unless somehow they are stinted just as the supply of money is stinted.

The fact of scarcity makes it necessary for us to economise, i.e., to make the most of what we have. We have constantly to be counting the cost, weighing up alternatives, and going without one thing so as to be able to buy more of another. Nominally it is money that we economise, for what we have to decide is whether to spend money on this or on that. What we are really doing, however, is to economise the things that money will buy. We try to buy, with our limited income, the collection of goods and services which gives us most satisfaction. We are faced with the fact that these goods and services are scarce, and we have to accommodate this scarcity as best we can to our wants and needs. Similarly, in earning money we have to husband our scarce time and energy in order to obtain as large a return as possible (in money *or* in amenities and personal satisfaction) for our efforts. On some men, of course, the pressure of scarcity and want bears harder than on others. On the millionaire, for example, the pressure is negligible; he can almost always neglect considerations of cost. But for others the necessity of making ends meet enforces constant self-denial.

**The Economic Problem.**—What is true of each of us is true also of society as a whole. There is an economic problem of making the *social* income go as far as possible. The goods produced and services rendered in any country in the course, say, of a year, are limited in amount and insufficient to maintain a standard of more than moderate comfort if equally distributed amongst the inhabitants of the country. The goods and services at the disposal of the country, in other words, are scarce in relation to the demand for them. There are very few things that can be provided free of charge, even in a rich country like Britain. We can make as much use as we like of public libraries and parks and roads. But we cannot help ourselves to books and motor-cars, much less to food and clothing. The more of one thing is offered to us, the less can we have of other things. We cannot have more of *all* simultaneously. If A is free, B will be all the dearer. The provision of free motor cars, for example, would lead to an expansion of the automobile industry and the transference to it of engineers, materials, and machinery from a host of other industries. Motor-cars would be more abundant; but other things would be scarcer. Only if we set a very high value on motor-cars (like the value which we set on good roads, or schooling, or health services) will we be prepared to face the cost of offering them free.

This balance between value and cost is forced on us wherever we are faced with a shortage of supplies relatively to our wants. The things which we value highly and which cost little to produce will be provided first and in large quantities. What costs a great deal and is of comparatively little value will not be produced at all. We have to

decide what commodities, and how much of each, to produce; and our decision will rest upon our estimates of cost and value. The decision is one that must be taken in every society, whether it be Russia or the United States, Italy or Melanesia. The way in which the decision is taken, and the kind of people who take it, are, of course, very different in different countries. The responsibility may rest with a bureaucracy or with the mass of "consumers." One country may have a State Planning Commission; another may rely on the laws of supply and demand. Whatever the economic system, the decision is one that cannot be avoided. There is an economic problem which has to be solved by dictatorships and democracies, "planned" and "unplanned" societies alike. Want and scarcity are universal, and so, too, is the problem of accommodating the one to the other.

In some countries the problem may be solved more satisfactorily than in others. But there is no question of one social system bringing plenty and another condemning us to scarcity. Man's wants are insatiable, and there would continue to be scarcity under any social system. If, for example, we all had twice as large an income as at present—an advance which could not be brought about immediately by any conceivable change in our social system—the annual income of the average British worker would still be under £600, and from this sum a large slice would be taken in taxation, and a further slice would have to be put aside as savings. Such an income would probably fall short of the aspirations of most people and could be reached only by exertions which would be decidedly irksome. The conflict between scarcity and want would continue to be felt.

Scarcity, like exchange, raises problems for the economist. He tries to formulate the principles on which our limited productive resources can be used to the fullest advantage. He studies how unemployment, for example—an obvious waste of labour power—can be reduced or eliminated; how the community's savings can be made to find their way into productive investments; how the land can be cultivated in the best interests of society. He studies, too, on what principles we should allocate resources between different industries so as to produce a maximum of all commodities in the right proportion of each; and how the output of commodities should be distributed between those who help to make them. These are problems which cannot be confined within the narrow bounds of pure economics. They extend into politics, ethics, and even religion. But we can get a better view of them from the heights of economic theory than from any other standpoint. Since this better view will still be coloured by our personal convictions, it will not of itself remove differences in outlook. But it will give us a wider perspective and open our eyes to the more remote implications of our problems.

### 3. Choice.

The use of money also implies choice. We have to choose between the many claims on our purse when we spend money, and between the many uses to which we might put our time and energy in earning

it. We cannot spend the same evening in the cinema and in the theatre. We must choose one form of entertainment or the other. We may have to choose, also, between spending an extra shilling or so on a seat and spending the same shilling later on cigarettes.

Our choice, of course, is not always made rationally. That is, we do not always weigh up carefully the possible ways in which we might spend our money. We are much more lighthearted and irrational in buying sweets, for example, than we are in renting a house. We buy, very often, impulsively or through habit or force of example. Or we may buy because our "sales resistance" has crumpled at the sounding of some advertiser's trumpet. It is irrational to pay more than is necessary for a thing, and yet hardly a day passes but we buy goods without asking their price, or cannot be bothered to look for cheaper brands. We do not take the trouble to find out where prices are lowest, or we take excessive trouble to save a trifling sum, like the wealthy man who walks to save a penny fare. We do not budget for so much on clothes, so much on amusements, so much on food, so much on our savings account, and so on, but spend haphazard so long as the money lasts. Or at least that is what large numbers of us do. Perhaps, however, the careful housewife—and the tradition amongst economists is to think of housewives as the persons who hold the purse-strings—is more rational in her buying The economic woman may be less of an abstraction than the economic man!

The way in which we make a choice is of great importance to the economist. For he cannot tell how much weight to place on the preferences expressed in the spending and earning of money until he knows how far these preferences are *rational* (i.e, based on full knowledge and formed after reflection) If, for instance, people persist in buying an expensive brand of cigarette it is important to know whether they buy it out of a liking for that particular brand or because they are ignorant of cheaper brands with the same flavour or because of snob-appeal in the advertisements Until the psychology of cigarette-smokers is explained to us, we cannot say whether the production and the sale of these high-priced cigarettes involve a social waste. If smokers are rational there *may* still be a waste (for instance, the price may be kept high by a monopoly) But if they are irrational, there is certainly a waste, they are paying more than they would if they were in possession of all the facts

In economics we begin by assuming that choice is rational The so-called "economic man" is simply one who is completely "rational" in satisfying his wants, and pays no regard to the interests of others. It is, of course, an abstraction from the facts to assume that men are self-interested and rational But to make this kind of abstraction is the only satisfactory procedure open to us. If we assume that people are self-interested and rational, we can predict how they will behave given a certain monetary inducement, and we can work out an analysis of action and reaction. For instance, if

similar goods are on sale at different prices, or similar jobs advertised at different rates of pay, we know that men will, other things being equal, purchase the cheaper goods, and apply for the better-paid job. If we could not make such generalisations, if men were quite irrational, then we should never " get anywhere " in economics. So we begin by assuming that choice is deliberate and rational, without, however, overlooking the part played by impulse, custom, and inertia. Later, we may study the psychology of choice more closely; analysing what shapes our expectations and desires, and sifting what is basic in our wants from what is superficial or conventional. But to begin with, we ignore these difficulties, take people's desires for granted, and assume that choice is rational.

In the economic system as we know it in Great Britain, choice rests largely with the individual. His preferences go to determine what is to be produced and what is not. Every penny spent on A is a vote in favour of the production of A; every refusal to buy B is a vote against the production of B. It is the free choice of individual consumers between the goods competing on the market that determines what industries can carry on at a profit. The industries that cannot show a profit are not carried on at all. Those that show excessive profits attract competition and expand until people's wants—as indicated by the price which they are prepared to pay—are more adequately met. That is, if competition is possible and effective. But if some commodity is monopolised, consumers may be powerless to get what they want (and will pay for) in the proper quantity. They show their readiness to cast votes for more of the commodity by offering high prices for it. But the election is disregarded. No one is willing to stand against the monopolist. So he is able to preserve an excessive scarcity by keeping people out of his line of business. He makes things scarcer than people want them to be and earns high profits by doing so.

Thus a country like ours does not deliberately decide what industries fit best with its advantages and needs and on what scale they should be carried on. The decisions that might otherwise rest with a central planning authority take shape instead in the market. One industry expands and another contracts as consumers alter their preferences and purchases. The scarce productive resources of the community are not rationed between the different industries by some Planning Commission. They flow into the channels lubricated by the expenditure of consumers.

But is it desirable that the individual should retain so much freedom of choice? What if consumers are irrational or incapable of judging between competing goods? Would it be better to appoint a State Planning Commission with power to decide what kind of goods should be produced and what kind of jobs workers should be encouraged to take up? Should each man's daily rations be assigned to him as the average man's daily work is at present? With whom should choice rest, and through what agencies is it best exercised? Here is another batch of problems for the economist.

**A More Precise Definition of Economics.**

We have seen that the use of money implies exchange, scarcity and choice, and that it is these things rather than money itself which are the starting-point of economics. Hence it is necessary to recast our preliminary definition of economics as the study of the part played by money in human affairs. We must find a more precise definition which goes to the roots of the matter. But which root? Some economists stress exchange, and others scarcity and choice. Is economics the study of man's behaviour in the act of exchanging, and of the problems raised by exchange? Or is it, as Professor Robbins argues, "the study of human behaviour as a relationship between ends and scarce means which have alternative uses"? This second definition starts from the fact that our wants are many and the means to their satisfaction limited. It emphasises the need for economy—the need to make the most of our scarce resources by choosing carefully between the wants which these resources may be used to satisfy. The definition might be re-worded to read: Economics is the study of the influence of scarcity on human conduct in circumstances where men have freedom of choice in allocating scarce resources between competing wants. Thus the definition runs in terms of scarcity and choice, in contrast to the first definition which runs in terms of exchange.

Now it is clear that scarcity is more fundamental than exchange. It is, in fact, as a result of our efforts to deal with scarcity (i.e. to economise) that exchange arises. We try to ration our limited means among the innumerable wants that compete for satisfaction and find that we can make our limited means go farther by striking bargains with our neighbours. We give what we have in relative abundance—muscle or brain, professional knowledge or organising ability—for what is comparatively scarce, what we could not do, or could not afford to do, ourselves. We sell our time and energies and spend our earnings on what others have laboured to produce. In doing so, we are offering the goods or services in which our talents show to greatest advantage (or least disadvantage) for the goods or services which others are specially fitted to produce.[1] We are supplementing our deficiencies—our imperfect versatility, for instance—out of the proficiencies of others. Not only are we able to draw on the skill of others—skill which we may not possess at all—but we are also able to give our whole energies to a single task—one to which, either through practice or natural bent, we are far more fitted than those who engage in it only intermittently. By exchanging, we are making our efforts go further towards meeting our wants. We are reducing the pressure of scarcity and achieving economy.

[1] This must not be taken to mean that everyone is engaged in just that occupation for which he is ideally suited—that would obviously be nonsense. What is meant is that people tend to specialise, *so far as their limited opportunities permit*, in the occupations which give fullest scope to their peculiar gifts and inclinations.

Exchange arises, therefore, in order to alleviate scarcity, and scarcity is necessarily implicit in exchange. But is it true that scarcity by itself gives rise to economic problems? Is there any room for the economist in a society where exchange has ceased to exist? Suppose, for example, that we lived in a Robinson Crusoe type of society in which everyone supplied his own requirements and was too busy "cultivating his garden" to exchange anything with his neighbours: or that society was organised on a Christmas Presents principle, every man working and giving away what he made, and living on the gifts that he received from others. Would these systems give rise to any specifically economic problems? The obvious answer is, no. There might be muddle and waste and bad accounting. But the problems of an exchange-system—for instance, those connected with tariffs, unemployment, taxation, and so on—would have no existence. In the Christmas Presents society, choice is, so to speak, vicarious. People have to content themselves with whatever others provide for them, and are free to devote their time to their private hobbies, unharassed by the problem of making ends meet. There is no economic problem.

Robinson Crusoe raises more difficult issues. Although he cannot exchange, he can exercise choice in conditions of scarcity. He illustrates how people set about satisfying their wants when exchange is impossible, and the study of his problems may be a useful preliminary to the study of exchange itself. But on his success in achieving economy, the economist, as economist, has little comment to offer. It may take an accountant or a philosopher to tell us when a man is getting the most out of his income, or making the most of his time. But it does not take an economist—who may be no more economical in this sense than other people. His job is to study economising once it takes on a *social* aspect: that is, once our attempts to economise impinge on the attempts made by other people. In our private lives we economise many things—words, patience, golf-clubs, and so on—that have no place in economics, because our acts of economy have no social repercussions. The problems of economy that face us are not always problems in economics.

To sum up in a compromise definition: *Economics is a social science studying how people attempt to accommodate scarcity to their wants and how these attempts interact through exchange.*

**Economics and Other Subjects.**

A precise definition like this immediately suggests that there is a clear line of division between economics and allied subjects. This is not so. Economics overlaps with many other subjects which study our wants and their satisfaction—psychology, ethics, advertising, and so on. The connection between these subjects may be illustrated from an example—cigarette-smoking again.

**Ethics.**

Let us start with ethics. Suppose that I want to smoke a cigarette in defiance of my parents' wishes, or in a crowded railway compart-

ment. I may pause to ask whether it is right for me to smoke. This is a question for ethics, which discusses the moral aspect of wants. If, however, I ask whether the whole practice of cigarette-smoking is wrong, I raise an economic as well as an ethical problem. I cannot decide whether smoking is good or bad until I know a good deal about the *consequences* of smoking, and to obtain this knowledge I may have to consult an economist. I may be afraid to condemn smoking because, for example, if everyone gave up smoking, there might be a big increase in unemployment. An economist would be able to assist me in judging what stress to lay on this danger.

Not only do ethical problems lead on to economic problems: the reverse is also true. It is hard to keep a discussion of any economic problem free from moral judgments of right and wrong. If I ask whether cigarettes sell for more than they cost to produce, and why that should be, I am sticking to "pure" economics. But immediately I ask whether it is *fair* to charge such high prices, or whether it is *right* for girls to be paid 8d. for each hour of drudgery in a cigarette factory while the managing director earns perhaps thirty times as much, or whether it is *in the public interest* to admit Virginian and Empire tobacco to the British market on the same terms—in fact, if I ask any of the questions I am likely to ask—I cannot be answered to my satisfaction unless I make plain what I consider "fair," or "right," or "in the public interest": that is, unless I lay down definite principles of fairness and justice. In the problems of the real world, moral and economic issues are hopelessly tangled up. We can unravel the threads, calling this moral, and that economic. But in the answer to the problem, they must be woven together. However reluctant economists may be to introduce the brittle thread of ethics (so often snapped by disagreement) into their analysis, they cannot offer the guidance which is so urgently sought of them unless they do. They can explain how the economic system works without putting the mantle of philosophy over the rather drab working clothes of economic science. But they cannot say how the system can be made to work *better*. They can offer light, but not fruit; and it is fruit for which most people turn to economics. Immediately the economist does venture to offer counsel—as is expected of him—he appears in the rôle of sheep in wolf's clothing, economist turned philosopher. This is a rôle which he must play consciously—not sheepishly, as if there were no wolf's clothing there!—if his conclusions are to command attention and respect.

**Politics.**

The same difficulty of demarcating the province of the economist is also raised by politics. Again there is apparently a clear line of division in subject-matter while in practice the line is often difficult to draw. Politics studies our wants, not for goods and services, but for things like equality, justice, and order. These are wants which are satisfied, not by reducing the pressure of scarcity, but by making people behave in certain ways. If we want order, for example, we

want power to impose standards of conduct on our fellows; we do not allow people to behave as they choose, but pass laws denying some rights and upholding others. Thus if smoking were forbidden by law, the justification of this provision would be a matter for the political philosopher. He might try to discover, for instance, whether, if the principle involved were extended to parallel situations, its acceptance would be likely to command general approval. Quite possibly he would find it hard to decide without some knowledge of economics. This would be even more essential if a proposal was made to tax smoking severely without prohibiting it. If, finally, the Chancellor of the Exchequer wanted advice about proposals to raise the duty on tobacco and to grant a subsidy to tobacco-growing, he would be likely to consult an economist rather than a political philosopher. In giving advice, the economist would again be filling a dual rôle, for the questions put to him would raise issues in politics as well as in economics.

How far-reaching this overlapping of economic and political problems might become can readily be appreciated if we suppose that the onus of deciding what commodities to produce, and how much of each, came to rest entirely with Parliament. The consumer would cease to vote for this industry or that by spending money on its products and would vote instead for the candidate who promised him more guns or butter or whatever he happened to want most. The elector would replace the consumer. A system of rationing would be in general use and exchange would continue only on a limited scale. Economic and political issues would be almost indistinguishable.

**Why Economists Disagree.**

The close connection between economics, ethics, and politics—and we shall meet with plenty of examples later—makes it easy to understand why economists disagree. Their recommendations must be an amalgam of science, philosophy and practical wisdom. Hence they are likely to disagree whenever the verdict of economic science is ambiguous or whenever they do not share the same political sympathies, the same conception of social justice, or the same intuition of what is practicable. The more the problem admits of scientific treatment, the more do economists approach to unanimity. But even in purely technical problems, into which ethics and politics do not enter, there is plenty of disagreement. It is far more difficult to discover laws governing human behaviour than scientific laws governing the behaviour of atoms.

**Economics and Other Subjects.**

The relationship of economics to the other subjects studying human wants is simpler to explain. If we ask, for example, why people want to smoke, we shall have to go to psychology, and probably also to history and physiology, for an answer. If we ask how cigarettes are produced, we shall be forced to study the technique by which tobacco is grown, blended, and packed, the technique of engineering, the technique of paper making, the technique of printing, and so on indefinitely through

the whole range of science. In economics we begin by taking wants and technique for granted and do not inquire either how the want arises or how it can be satisfied. Our questions are in terms of cost and value, not of motive or technique. But it is impossible to ignore either of these. Take motives, for example. The broad assumptions on which economics is built up are essentially assumptions about motives, and these assumptions may, if we neglect psychology, lose touch with the facts. We cannot lightly assume, for instance, that the general desire of business men is to maximise their profits and that this desire does not conflict seriously with more ultimate desires like those for power or social distinction or for creative activity and a rôle in life. A business may stop short of its most profitable size or expand beyond it because the owner does not think exclusively in terms of maximising his profits. The more complex men's motives are, the more does a knowledge of psychology become indispensable to the economist. Similarly with technique. There may be nothing in the description of, say, a coal mine to call forth the special talents of the economist. But if he takes the trouble to see a mine at first hand he may cease to make foolish statements that bring his profession into disrepute. That is, he should know what he is talking about. Before discussing speculation, for example, he should know something of stock and commodity exchanges; before discussing the price of cloth, he should be familiar with the chain of processes that lie between the growing of raw cotton and the finishing and dyeing of cotton fabrics. He should know how things work.

**The Value of Economic Studies.**

Clearly, economics is a very complicated subject—so complicated that many students ask whether it is worth their while to learn a little of it when there is so much that they will never have time to learn. The easiest line of reply is that we do not fly from the study of mathematics in terror of the differential calculus. More convincing arguments can also be put forward. For some people, the value of economics—even in fairly small doses—lies in disciplining them in sustained thinking about social problems. The student who is accustomed to a slipshod dogmatism, and is unaccustomed to tracing the consequences of admitted facts through chains of reasoning, stands to gain enormously from the discipline which the study of economics imposes on him. He comes to see what it is that he is assuming and becomes alive to the wider implications of those assumptions. He is forced to state his argument quite unambiguously, so that verbal misunderstandings are cleared away. He learns to be more impartial and look at all sides of the question; to generalise properly by distinguishing what is relevant from what is irrelevant; and to pay more attention to the average, inconspicuous case (as revealed, for instance, by statistics) and less to the few striking instances under his nose. For others, the value of economics lies mainly in shaking them out of an unreasonable complacency in their political philosophy. They may be all for changes in the wage-system or in the money-system

which are either demonstrably impracticable, or practicable only if accompanied by further changes in which they are not prepared to acquiesce. Or they may be for no changes at all, without having weighed the case for and against. They may be roused to think for themselves on encountering the wide range of points of view which economists express, and accept less readily the dogma and propaganda purveyed so liberally by all the agencies which control public opinion.

Above all, economics is of value in allowing us to judge and frame policies in the light of full knowledge of how the economic system works. Everyone, willy-nilly, is an economist, for everyone has his view of how economic forces work, and what, if anything, should be done to control them. We all at some time or another maintain that workers should be paid a "living" wage, or that there is a deficiency of purchasing power, or that rents are too high, or some such proposition. These are propositions which one must be an economist of sorts to debate. But, of course, "economists of sorts" are not trained economists. The trained economist discusses propositions of this kind with the help of a whole apparatus of thought—what is called "economic theory"—which has the same advantage over untrained common sense that medical science has over popular medicine: with this reservation—that the ordinary man has a much sturdier disbelief in the pretensions of economists than in the pretensions of doctors. He makes no claim to understand or dispense with the higher calculus; but economic theory, chiefly because of an exaggerated impression of the failure of economists to agree, is all too frequently dispensed with The result is that elementary misconceptions flourish everywhere: in the conduct of business, in moral judgments, and in public policy. Economics is to these misconceptions what chemistry is to alchemy, or astronomy to astrology. It does not furnish a panacea or a creed. But it does serve the negative purpose of discrediting the crank and the charlatan; and it inculcates a technique of thinking without which a sound judgment of affairs is hard to come by.

One final point. It must not be supposed that economists are blind to everything but economy and efficiency. They may, privately, have a great liking for uneconomical and inefficient ways of doing things. But they assume that most people would prefer to work less hard or in more congenial occupations if they suffered no loss of income in doing so; that is, they assume that an economy of human effort is, in present circumstances, desirable, and that if there is some more efficient way by which the work of the world could be done, then it would be a pity not to try to discover it. They are interested mainly in *money*-incomes and *money*-costs—in making money-incomes high and money-costs low. But they do not suppose that every rise in incomes and every cheapening of goods and services is an advance in human welfare. For they know that the things that money will buy are not what men most desire; and that cheapness in money-costs may conceal dearness in flesh and blood, in social amenities, in human needs and enthusiasms.

Economists, then, do not regard money and the things that money will buy with a special veneration. But they know, more vividly than others, how enormously monetary considerations influence human conduct. They know that those who affect to despise money are generally well provided with it. They know that men tend to believe what it is in their interest to believe and that men's interests are inseparably linked with their bread and butter and the way it is earned. They know how unemployment and poverty can change men's minds as well as their bodies: how the course of our lives depends on the income and occupation of our parents: how the desire for power and mastery can be turned to acquisitiveness and greed. Knowing these things, economists are in no danger of underrating the economic factor in human affairs. But they can also avoid exaggeration. They have learnt something of the plasticity of human motives and ambitions and know that the hope of gain and the fear of unemployment have not always been the main incentives (and will not always be the most effective ones) for contriving that the work of society is done. Money and jobs are not the only human obsessions. It happens that the rôle which a man plays in life has come to depend more and more on the economic function which he performs and the social standing which his income obtains for him. But the time may come when prestige attaches to other and more important things than a man's income or his job, when the tiresome necessity of earning a livelihood has become less pressing, and qualities of spirit have come into their own. Such a time economists will hail with delight; or perhaps the study of their subject will have become superfluous, and economists will be extinct.

## CHAPTER 2

## SOME PRELIMINARY CONCEPTIONS

THE economic problem is one of accommodating scarcity to human want: in other words, of overcoming scarcity and poverty. Our success in solving the problem is measured by the amount of wealth which we produce, and the fairness with which we distribute it. If we start from the problem, we study scarcity; if we start from the solution, we study wealth. In the first chapter we took the first point of view; in this chapter we shall take the second, and start from the production and distribution of wealth.

**Wealth.**

Wealth, in economics, generally means a stock of goods possessing a money value. This implies three things. First, before anything can acquire a money value it must possess utility or the power to satisfy a human want; what has no utility is obviously valueless. Second, nothing can have value unless it is limited in supply; if there is more than enough for everyone, no one will be willing to pay for a

share. *Thirdly, goods can be given a money value only if they are* marketable; what cannot be bought or sold cannot be valued. Wealth, therefore, consists of the stock of marketable goods, which, in his struggle to overcome scarcity, man has built up for the satisfaction of his wants.

(1) Wealth must possess utility. Utility, however, measures the strength of our desire for a commodity—its "*desiredness*"—not its desirability. In deciding what is wealth and what is not, therefore, we pass no moral judgment on the wants which are satisfied. To those who buy them, drugs are wealth; on the other hand, sunshine, being free, is wealth to no one. Wealth satisfies the wants of saints and sinners alike; the fact that it is increasing is no evidence that there are more saints and fewer sinners, or that the desires which are being fostered and gratified are wholesome and healthy rather than vulgar and vicious.

(2) Wealth does not include everything that possesses utility. It includes only those goods which are limited in supply. Other goods, like air, water, and the free gifts of nature, are free for everyone to enjoy and, therefore, cannot be valued in terms of money. Goods of this kind, however important for human welfare, do not give rise to any economic problems, since they cost nothing and need not be economised. Goods which are free in some circumstances, however, may become relatively scarce in other circumstances, and so pass from the category of free goods into the category of wealth. Air at the seaside is a free good; but an adequate supply of air in a large cinema calls for a system of ventilation, the installation of which may be very costly. Similarly, water in time of drought and light at night-time cease to be free and come to acquire a money value. The supply falls short of our potential demand for them, and a price has to be charged so as to limit demand to the available supply and to compensate those who are trying to make good the shortage in supply. Once a price is charged, economic problems emerge, and the goods come under the heading of economic goods, or wealth.

(3) Wealth includes only those goods which are marketable. A man's beard, for example, while it no doubt satisfies a very human want and is limited in supply, is not wealth because its ownership cannot be transferred. Similarly, personal qualities such as health, good taste, capacity for enjoyment, natural talent, and so on, form no part of wealth since they cannot be separated from their owner and sold. This is true even of qualities which are of great advantage in business, e.g. a man's skill, or credit, or his organising ability. These are creative of wealth: the services of a skilled man are more valuable than those of an unskilled man. Personal qualities are also alternatives to wealth; we may have to choose between spending money so as to improve our skill or education, or for the purchase of a house or motor-car. But personal qualities cannot, except under slavery,[1]

[1] A further exception might be made in favour of artists under contract to a company. The actual qualities, of course, are not transferred, but the company acquires control over their use—a "property" in them.

be transferred from one man to another and acquire a market value, and they must, therefore, be excluded from wealth.

Although wealth must be marketable, it is not necessary that it should be marketed. There is, for example, a great deal of property in collective ownership—e.g. roads, bridges, libraries, schools, etc.—in which there is no market but which is unmistakably wealth. It would be clearly be absurd to exclude roads from wealth and at the same time to include railways, which serve a similar purpose. Since roads are provided free, however, we have much more difficulty in estimating their money value than in estimating the value of the railway system. Railways can be valued by estimating what profits they are likely to earn and capitalising these profits (i.e., converting expected profits to their capital value by mutiplying them by an appropriate number). But roads, since they yield no revenue, cannot be valued in this way. They are valued, therefore, on the basis of their cost of production, although this may be far in excess of, or far below, their true value.

**Wealth and Welfare.**

In everyday speech, wealth is not always carefully distinguished from welfare. We say, for example, that " health is wealth " without being quite clear whether we are referring to the higher income (wealth) which a healthy man can earn, or to the sense of well-being and enjoyment (welfare) which he experiences. It is plain, however, that the two things are quite different from one another. Wealth is a stock of goods, welfare a state of mind. Wealth may be *creative* of welfare by satisfying our wants ; but it cannot *be* welfare.

The accumulation of wealth is not *necessarily* creative of welfare. The wants which are satisfied may be evil or degraded. Each increase in wealth may whet the appetite for more, until civilised pursuits are crowded out by the scramble for riches. Or we may give our time and energies to the production of wealth at the cost of health, leisure, or peace of mind. What is true of our personal welfare is true also of social welfare. Riches bring no more assurance of happiness or well-being to communities than to persons. An increase in a country's wealth, for example, may excite envy abroad or widen social cleavages at home. Roads and houses may be built while the human beings who will use them are badly fed, ill-educated, or over-worked. The countryside may be disfigured by a growth of ugly bungalows. In these and a hundred other ways, the wealth of a country may increase at the expense of its welfare. Nevertheless, an increase in wealth will, *as a rule*, be a real contribution to human welfare, while a decrease will be a real social loss. It is upon this assumption that economists proceed.

**Wealth from the Individual and Social Points of View.**

What is wealth to the individual is not necessarily wealth to society. The racketeer who has built up a useful " connection " with local shopkeepers is the owner of a valuable business, the " goodwill " of which (were there a market in such businesses) might run into thousands

*of pounds. But he adds nothing to wealth from the social point of view. He is paid to satisfy a want (for "protection") which is* entirely artificial, and which would cease to exist if he went out of business. Out-and-out racketeering, fortunately, is very rare, but there are some businesses, of which it is true, just as it is true of genuine racketeering, that if they were to be closed down, there would be no great loss to society. A quack may induce the public to buy medicines which are either worthless or positively harmful. A coalmine near a residential district may ruin the amenities of the district without being required to compensate neighbouring householders. Or, if a number of competing firms are all working under capacity, the closing down of one of them may expand the market of the survivors and increase their efficiency by allowing them to produce nearer to capacity.

In all these examples, the value of a business to society is less than its value to its owner. Sometimes, however, the opposite may happen. If the railway system, for example, were out of order, all the industries that use it would be disorganised and the loss to society would be greatly in excess of the loss to the railway companies. The same would be true of the Post Office and of electric supply companies. Power, transport, and communications are the nerves of modern industry and if they stop functioning, industry is immediately paralysed.

Thus a business may contribute either less or more to the wealth of society than it does to the wealth of its owner. Whenever such a divergence between value to the individual and value to the community arises, private advantage and the social interest come into conflict with each other. A business which is making a profit does not close down simply because society would gain from its suppression, nor does a business continue to operate at a loss even when there is a balance of gain to society. We shall come across many examples of this conflict of interest between private and social advantage in succeeding chapters.

**Wealth and Titles to Wealth.**

When a private person is calculating his wealth, he includes not only goods in his possession (e.g., his house) but various documents such as stocks and shares, I.O.U's, pound notes, etc. From the social point of view, however, these documents are not wealth. They are titles to wealth or expressions of debt. Stocks and shares, for example, are titles to property controlled and operated by a joint-stock company. It is this property which we include in social wealth and not the documents in evidence of ownership. Similarly, if someone lends money and is given a bond as proof of the debt, any property built out of the borrowings is included in social wealth but not the debt itself. If we did not do this, and counted in both the debt and the asset built out of the proceeds, we should clearly have double-counting. In addition, the inclusion of debts in social wealth would yield the paradoxical result that a country with a very large National Debt (incurred, for example, to pay for some past war) would be wealthier than a neighbouring country with no National Debt.

**The National Wealth.**

The social wealth of a country is generally referred to as its National Wealth or National Capital.[1] This may be defined as the aggregate value, at any moment, of the stock of assets, wherever situated, in the possession of its inhabitants. Wealth inside the country belonging to foreigners is thus excluded, while the value of the country's investments abroad is included. No account is taken of debts due from one citizen to another, for these debts will appear as credits in one man's wealth and as debits in another's, and they will therefore cancel out when we take the aggregate of private wealth. All property in public ownership is added in, while a deduction is made for any debts incurred by public authorities (e.g., the National Debt).

Before the war, the National Wealth of Great Britain was probably rather in excess of £20,000 millions. Estimates made by the late Lord Stamp[2] for Great Britain and Northern Ireland show an increase as follows :—

| | £000 millions |
|---|---|
| 1914 .. .. .. .. .. | 13·7 |
| 1928 .. .. .. .. .. | 18·2 |
| 1935 .. .. .. .. .. | 21·5 |

Mr. Campion[3] reaches rather lower totals, but the difference (especially in view of his exclusion of Northern Ireland) is relatively trifling. Both he and Lord Stamp exclude the value of British roads and armaments.

In comparing the estimates for the years given above, we must bear in mind that changes in the real National Wealth may be masked or exaggerated by changes in the level of prices. The money counters in which the National Wealth is reckoned have not a constant value ; so that when the value of money changes, the National Wealth in terms of money will change also, although in real terms it may have remained constant. We must not assume too readily, therefore, that we had more wealth in real terms in 1935 than in 1914. In actual fact, prices rose by about 40 per cent. between these dates, so that the real increase in wealth was probably a good deal less than the apparent increase.

The total of just over £20,000 millions for pre-war years was made up roughly as follows :—

| | £000 millions |
|---|---|
| Dwelling-houses .. .. .. .. .. | 5·0 |
| Furniture, motor-cars and other durable goods | 1·0 |
| Overseas investments .. .. .. .. | 4·0 |
| Agricultural land and farmers' capital.. .. | 1·0 |
| Railways .. .. .. .. .. .. | 0·7 |
| Industrial and commercial buildings and plant | 7·0 |
| Property owned by public bodies .. .. | 2·3 |
| | 21·0 |

[1] For the relationship between wealth and capital, see page 35.
[2] Lord Stamp : " The National Capital and other Statistical Studies."
[3] H. Campion : " Public and Private Property in Great Britain."

It will be observed that only about a third of our wealth took the form of land, buildings and equipment owned by industrial and commercial undertakings. A large part consisted of overseas investments, and a still larger part of durable consumers' goods such as houses and furniture.

**Wealth and Income.**

Wealth is sometimes used in the sense of capital (i.e., a stock of goods at some point in time) and sometimes in the sense of income (i.e., a flow of services over a period of time). It is best to use wealth in the first sense only. Wealth and income can then be sharply distinguished from one another. Wealth consists of stored up facilities for the satisfaction of human wants; income consists of the flow of services yielded by wealth. For example, a house is wealth, while the shelter which it affords is income. Coal is wealth while it remains in the scuttle; it becomes income when we put it on the fire. Income is not confined, however, to the value of the services yielded by *goods* when we use or consume them. It includes also the value of services (like the waiter's, the 'bus conductor's, or the lecturer's) which are rendered by persons. Normally the services which we render in our work are congealed in goods—as the work of a carpenter, for instance, is congealed in tables and chairs; later, when these goods are consumed or used up, they render a service. First we produce wealth, then the wealth yields income as it is consumed. Wealth, therefore, is a link between our efforts and the wants which these efforts satisfy. Sometimes, however, wants can be satisfied directly without wealth coming into existence at all. The services of a dentist, for example, yield nothing that can justifiably be called wealth; but the dentist contributes just as effectively to the satisfaction of our wants as the carpenter or the butcher, and his contribution is just as entitled to be included in income. Income, therefore, includes the value of services rendered as well as the value of goods consumed; indeed, as we shall see, the value of goods consumed is ultimately reducible to the value of the services which they embody.

**The National Income.**

We can measure the National Income of a country in any one of three ways: by adding up everybody's income; by adding up everybody's output—the value of the goods that we are paid to produce and the services that we render; or by adding up the value of all the things that people buy and adding in their savings.

If these different ways of reckoning the National Income are to give the same answer, they must obviously proceed on a uniform definition of income. The goods and services that enter into income must be taken to include only those that are in fact customarily exchanged for money. We can make no provision for services which we perform free of charge and do not normally count as part of our income (brushing our own boots, cooking the family dinner, etc.). Similarly, we must exclude the benefits which we derive from

furniture, clothes, and so on, since it is impossible to put a money valuation on them.

Estimation of the National Income is complicated by the part played by the State, which has a revenue and expenditure of its own. If we could assume that all taxes were levied as direct taxes on income, and that the proceeds were used exclusively to provide public services, we could neglect the part played by the State and regard its activities as on the same level as those of a gigantic club to which we had to pay an annual subscription. But, in fact, the State is not simply a club. It is also an agency for transferring money from one set of people, who pay taxes or make insurance contributions, to another set of people who receive benefit or pensions. These "transfer payments" figure in the incomes of both sets of people. But since the same income cannot be spent twice, it should not be counted twice. In estimating the National Income, therefore, we must deduct from the total of individual incomes any sums received from the State, in benefit or pension. Payments made by the State for current services (e.g., to the Police Force) add to the size of the National Income by the value of those services, and no deduction, therefore, should be made for such payments.

It is usual to exclude from the National Income interest payments on War Loan. The money is paid out of the income of taxpayers in recognition of a service performed long ago, not in payment for a service done now. The assets created out of War Loan have ceased to exist and yield no present output to swell the size of the National Income. It seems natural, therefore, to treat interest on War Loan as a transfer payment which allows the taxpayer the satisfaction of honouring a pledge, but hardly forms part of his expenditure on goods and services. Observe, however, that if the Government, instead of defending the country from its enemies, had built a chain of dog-racing tracks by loan-expenditure, it would be able to make interest payments on the loan, and these payments, since they would correspond to a continuing service, would go to swell the National Income.

A further problem arises in connection with indirect taxation (i.e., taxation of goods and services). The price paid by the public for tobacco, for example, contains an element of tax which is used to finance some service provided by the State. If we count in both the value of the tobacco at market price and the value of the service—say, roadsweeping—which is met from the tax on tobacco, we are double-counting. We ought, therefore, to measure the purchases of the public at the prices that would rule in the absence of indirect taxes (or of subsidies).[1] What this comes to is that we should deduct from the aggregate of private incomes the indirect taxes (including rates) levied by the Government, and add in any State subsidies designed to keep down the cost of living.

[1] This is generally referred to as measurement "at factor cost."

Estimates of the National Income of Great Britain have been made by Professor Bowley, Mr. Colin Clark and others. The latest and most authoritative estimates are those issued in the White Papers published annually by the Government in connection with the Budget.[1] These estimates, using the definition of National Income employed in the White Paper of 1944, are :—

| | £ millions | | | | | £ millions |
|---|---|---|---|---|---|---|
| 1938 | 4,604 | .. | .. | .. | 1941 | 6,885 |
| 1939 | 4,968 | .. | .. | .. | 1942 | 7,604 |
| 1940 | 5,945 | .. | .. | .. | 1943 | 8,172 |

These figures can be analysed in a variety of ways : to show how incomes were earned and on what they were spent ; to show the distribution of income between classes ; to measure the comparative prosperity of this and other countries ; to provide an index of economic activity ; to assist the Government in planning its Budget ; and for many other purposes. No figures are so fundamental to an understanding of our economic system.[2]

The simplest deduction to make from the estimates given above is that, before the war, income per head of population averaged only £100. Income per occupied person averaged about £200, and per family about £300. But the average is somewhat deceptive. Probably half of all income-receivers in 1938 earned less than 50s. a week. In an average working-class district in Scotland nearly a quarter of the *families* had a joint income of under 40s. a week. Only about a fifth (the proportion varied greatly from district to district) had a joint income of over £5.

The greatest care is necessary in interpreting an increase in the National Income too hastily as a sign of improvement. The fact that the National Income was higher in 1941 than in 1938 did not imply that we were better off and that we could count on an improved standard of living after the war. Prices were over 20 per cent. higher and in real terms, therefore, the rise was less than half as great as in terms of money. Services previously rendered free at home had been taken over by canteens, restaurants, laundries, etc., and entered up at selling value. Part of the increase was at the expense of leisure (there was more overtime) ; part at the expense of health (we worked harder) ; part at the expense of personal freedom (more people worked for a wage or salary instead of on their own) ; and part at the expense of enjoyment (many people had to work in surroundings which they disliked).

[1] "An Analysis of the Sources of War Finance and an Estimate of the National Income and Expenditure in 1938 and 1940" (Cmd. 6261). For later figures see the similar papers for 1942 (Cmd. 6347), 1943 (Cmd. 6438), and 1944 (Cmd. 6520).

[2] For a full analysis, the reader should consult Professor J. R. Hicks: "The Social Framework."

**Production.**

Production is the making of goods for sale or the rendering of paid services. The " making " of goods, however, must not be interpreted too literally to mean the creation of some material object. We do not so much make things as make them more serviceable. We create, not material objects, but utility. When, for example, a man makes a pair of boots, he rearranges pieces of leather from a *form* in which they are not wanted into a form in which they are wanted. When someone transports the boots from factory to warehouse, he is taking them from a *place* where they are not wanted to a place where they are wanted. When someone stores the boots in his warehouse, he is holding them over from a *time* when retailers do not want them to a time when they do. There is no essential difference between these activities. They all create utility, either of form, place, or time, and so improve the facilities for the satisfaction of our wants ; that is, they are all productive.

Many people take a much narrower view of production. Adam Smith, for example, dismissed as unproductive " some of the gravest and most important, and some of the most frivolous professions : churchmen, lawyers, physicians, men of letters of all kinds ; players, buffoons, musicians, opera-singers, opera-dancers, etc.," on the grounds that " the work of all of them perishes in the very instant of its production." If we take " productive " to mean " productive of wealth," this is sound enough. But in economics " productive " almost invariably means " productive of income."[1] In this sense, the professions are clearly productive. They create nothing tangible, but they provide us with services of all kinds in just the same way as manual workers provide us with the fruits of their labours. There is no difference in principle between the making of a violin, and the playing of it in a concert-hall by Kreisler.

Sometimes people speak as though only those occupations were productive on which we depend for the bare necessities of life. On this view, shopkeepers, cinema stars, and university professors rank below farm labourers and bricklayers. Now whatever we may think of this ranking—and there is much to be said for it—it has nothing to do with productivity in the economic sense. It is an abuse of language to call an occupation " productive " when what we really mean is " indispensable " ; and it is the language of abuse to call an occupation " unproductive " when what we really mean is " sinful in the eyes of true believers "—believers, that is, in the simple life and the virtues of manual labour. In the same way, when people condemn the building of warships, or grouse-shooting, or distilling as " unproductive," because they would like to see them put a stop to, they are using the term " unproductive " as a mere label of disapproval. It is better to avoid this by using the term " undesirable " (which conveys our real

[1] So much so, that even when economists do speak of the production of wealth, they are almost certainly using " wealth " when what they really mean is " income." For an example, see the first paragraph of this chapter.

meaning), reserving " unproductive " for work which is of no service to any one.

Production, to sum up, is simply the creation of utility : that is, of the power to satisfy human wants. Consumption is the using up of utility when we come to satisfy our wants. When we produce goods, we build up a store of wealth upon which we can draw. But, strictly speaking, we do not produce or consume goods at all. We render and are rendered services, some of which are consumed on the spot, while others are congealed in bread and butter, collars and ties, ships, houses, and so on, and are used up in periods varying from a few minutes to hundreds of years. Our efforts are productive if they are of service to some one. If they are not, then we may fairly call them unproductive.

**The Factors of Production.**

Production is undertaken by what are called by economists the Factors (or Agents) of Production. These are classified under three broad headings into Labour, Land and Capital. Labour represents the human element in production—not just muscular effort (which can be replaced by machinery), but also skill, enterprise, organising ability, and above all, judgment. Land and capital are the external facilities which assist labour in production. Capital, in the sense of concrete capital or property, is the stock of assets that has been built up by human effort and thrift in the past—factories, machinery, and so on ; in the sense of loan- or money-capital, it consists of money that is available for investment in new concrete-capital.[1] Land, unlike capital, exists independently of human effort : it consists of those free gifts of nature which are in human ownership and under human control (e.g., mineral deposits, agricultural land, fisheries, etc.).

Production can be carried on only with the co-operation of at least a minimum of all three factors. Land and capital can produce nothing unaided by labour—even if it is only the labour of reaping. Labour can produce nothing without *some* capital—even if it is only the capital out of which it maintains itself

**The Firm and the Industry.**

Production is organised in modern industry in more or less independent units called " firms." The firm is the unit of control over the factors of production ; it hires labour and land, purchases raw materials, borrows or puts up the necessary capital, organises the process of manufacture, and markets the finished product. Sometimes the firm is a one-man show—a costermonger or flower-seller ; sometimes it is a gigantic enterprise like Imperial Chemical Industries, Ltd., with separate establishments all over the country and all over the world. Sometimes, too, it is hard to say where one firm stops and

[1] An explanation of the connection between concrete-capital and loan-capital is postponed until the next chapter.

another begins—e.g., when the owners come to an understanding with one another and behave exactly like two branches of a single firm.

An industry consists of a number of competing firms. The firms may be producing the same commodity (e.g., coal), or working on the same materials (e.g., brass), or using the same process (e.g., building). There are thus three quite different bases on which we may group firms together into industries. Naturally this leads to difficulties of classification. Shipbuilding, for example, will be included in the steel industry if we use *materials*, and with building if we use *processes*, as a basis of classification. In practice, we rank shipbuilding as a separate industry, partly because it is big enough to justify the title, partly because its *product* is one in which many firms specialise. The crucial test is the way in which competition is delimited; and it is clear that shipbuilding firms are in active competition with one another, but are isolated from the competition of other firms which use steel or which are engaged in building.

It is exceptional to find all the firms in an industry confining themselves to the same range of products or processes or materials. Firms do not specialise along similar lines, but often compete in several different industries at the same time. A shipbuilding firm, for example, may engage in coal-mining and steel-making. A cotton firm may manufacture artificial silk products. Railways may own hotels, golf-courses, and Atlantic liners. It is in a few industries only—Grain Milling, Ice, Tobacco, Match, Wallpaper, and Incandescent Mantles—that firms confine themselves within the boundaries of their trade and are completely segregated from other industries.[1] Again, even when the limits of an industry are clearly defined, there may be little in common between two firms within it. The Leyland Motor Company, which manufactures commercial vehicles, no more competes with the Riley Motor Company, which manufactures private cars, than the London Midland and Scottish Railway competes with the Raleigh Cycle Company. Each firm within an industry has its own peculiar style or quality of product, or its own peculiar methods of manufacture, or uses its own peculiar materials. Sometimes these peculiarities are trifling (e.g., in cotton spinning), but in most industries they are of considerable and increasing importance. They make it very difficult to draw a clear line of division between firms which are in direct competition with one another (e.g., two coal-mining firms in adjacent areas), and firms between which competition is rather remote (e.g., a coal-mining firm and an oil company). At best, the classification of firms into industries is a rough one.

**The Market.**

A firm requires to strike a bargain both with the factors of production which it employs and with the purchasers of its products. For example, it has to bargain with labour about the wages which it will pay, and with its customers about the price which it will charge.

[1] Final Report on the Fourth Census of Production (1930), page 61.

These bargains are driven in what is called "the market." The market, it should be clearly understood, is not a place—although bargains are still driven about wages at hiring fairs, and about prices on wheat, rubber, and other exchanges. The market, in economics, is simply the network of dealings in any factor or product between buyers and sellers. These dealings may be regular and organised, or they may, as in the market in second-hand violins, be spasmodic and unsystematic.

It is round the market in this sense that the whole economic system revolves. Production is ruled by market requirements: if goods cannot be sold in the market they will cease to be produced. Men's incomes are governed by what "the market" will offer for their services and for the hire of their capital. The vast impersonal force of the market shapes the environment in which we earn our livelihood: we are free only within limits set by the market. At times, when we are in the grip of unemployment, the limits seem narrow and irksome; at other times, when we draw on the bounty of distant lands for the everyday necessities of life, the limits seem singularly wide and generous. It is the business of the economist to throw light on these limits and to suggest how they may be thrust further back by a proper organisation of our resources.

## CHAPTER 3

## THE EXISTING RESOURCES: THE FACTORS OR AGENTS OF PRODUCTION

THE factors of production, as we have seen, are classified under three broad headings—Labour, Land, and Capital. This classification is convenient but rough. No two workers, for example, have exactly the same skill, or the same versatility, or the same attitude towards their work. The number of occupational groups alone runs into thousands. Thus when we refer to "Labour"—or to either of the other two factors—we are speaking, not of a block of resources each unit of which is interchangeable with every other unit, but of a collection of units each of which has its own peculiar idiosyncrasies.

### 1. LABOUR.

Labour is the most important of the factors of production. If the other two factors are underemployed or overdriven, it is only the profits of their owners that suffer. But if labour is unemployed or sweated, it means degradation and poverty for men and women, and a lowering of the quality of social life.

The main characteristics of labour are intelligence and judgment. This may seem surprising to those who think of labour as toil and exertion. But toil and exertion are far more characteristic of machinery, which can work on unremittingly for twenty-four hours

a day. Even in agricultural communities it is horses and oxen, and not human brawn that supply most of the necessary power. Work that does not call for the exercise of intelligence, and is purely mechanical, can (as the word "mechanical" implies) be taken over by machinery. It may happen, of course, that the machinery is costly to make and operate, or that muscle-power is abundant and cheap, so that men continue to be used for what is literally inhuman work. But the more machinery is improved, and the scarcer brawn becomes relatively to skill and judgment, the more will toil-saving inventions be introduced. Muscle-power and machinery are in direct competition with one another and the one can replace the other. But the work of the human mind cannot be replaced. There are devices which, as we shall see, make a little intelligence and judgment go a long way in modern industry. But there are none which eliminate them completely from business any more than from the rest of life.

In saying all this, we are not assuming that the introduction of machinery is an unqualified good, or that it makes work somehow more "human." A man's muscles and imagination, to which machinery frequently gives little scope, are as much a part of him as his intelligence and nerves. When the machine sets the tempo and not the worker, there can be little opportunity for initiative and craftsmanship ; and these are things the value of which is not exhausted within working hours. They enrich a man's personality in all his doings, and personality is a far more important product of industry than the goods and services which it is the humble duty of economists to discuss.

If it is judgment rather than toil that is labour's main characteristic, we can draw no sharp line of distinction between the labour of, say, a 'bus driver and the labour of the director of an omnibus company. There may not be the same technique of judgment in both jobs. The problems of negotiating a street corner are very different from the problems of negotiating for the purchase of a fleet of new 'buses. But both require intelligence and judgment. If the skills differ, the ingredients do not. For there can be just as much purposive control, planning, and adaptation of means to ends in driving a 'bus as in managing a company.

**Organisation as a Factor.**—It is for this reason that we must reject the suggestion that is sometimes made, that organisation—or, alternatively, enterprise—should be treated as a separate factor of production. There may be a number of people who call themselves directors or organisers and have rather more organising ability or rather more scope for organising than other workers have. But they are not a class apart. It may not be everyone who organises *men* in industry, with power to give orders under threat of dismissal. But everyone, be he office boy, labourer, foreman, or works manager, has to organise either men or things. Moreover each worker—and not simply the manager—will be paid more if his work involves the responsibility of organising. The greater the skill and judgment he

has to show, the more he can expect to earn. Some people draw higher incomes than others for their organising; but everyone who is not doing completely mechanical work draws *some* income for what is really organising. Labour is a blend of toil and organising.

**Enterprise as a Factor.**—Enterprise, or risk-taking, is clearly not a factor of production in the sense in which Labour, Land, and Capital are factors of production. It is not part of our resources. An enterprising disposition, no doubt, is. But then so are all sorts of other dispositions. If the pioneer and the man who will "take a chance" are valuable to society, so are the men who like dull and monotonous jobs. It takes both sorts to make the wheels of industry go round. No one suggests, however, that routine-liking, or solitude-liking, or speed-liking are separate factors of production. Why, then, do some economists single out Enterprise, and class it, as a separate factor, with Labour, Land and Capital?

The answer is that their classification is based, not upon the resources which are used in production, but upon the rewards which are earned by the factors. Corresponding to Labour, Land, and Capital, we have Wages, Rent, and Interest. What more natural than to add Enterprise as the factor corresponding to, and earning, Profit? Undoubtedly this is, for many purposes, a convenient classification. But it has one great drawback. It leads us, because of our constant association of profit with the single factor, capital, to associate enterprise in turn with this one factor. But capital is not the only factor that is exposed to risks and is paid for its enterprise. Labour runs risks (e.g., of unemployment or of slow promotion) which are often of far greater importance to society than the risks run by capital. The steeplejack and the coal-miner may not risk their capital: but they risk their lives. There are risky occupations as well as risky undertakings; and part of the wages paid in these occupations should be classed as a return to enterprise, in just the same way as part of the return to investors in risky undertakings is classed as a return to enterprise.[1] The same is true of land put to uses in which the return cannot be foreseen. There is an element of profit in wages and in rent no less than in the earnings of capital. The worker who leaves a safe job in search of another, the farmer who tries out a new kind of seed, and the landlord who prospects for minerals on his land are all taking risks and shouldering the possibility of loss in the expectation of ultimate gain.

## 2. Land.

Land as a factor of production has lost much of its former importance. In the days before the Industrial Revolution, perhaps two-thirds of the workers of Britain were employed in agriculture, against a mere twentieth at present. The owners of agricultural land drew

[1] This must not be taken to mean that in practice wages are related at all closely to the hazards run by workers.

in rent over a fifth of the total income of the country, whereas nowadays the proportion is barely one per cent. A corresponding change has come over economic theory. The discussions of rent which fill the pages of Adam Smith and John Stuart Mill have dwindled sadly in the modern textbook. Similarly in politics. The disinherited of earlier centuries agitated against the landlord, as the representative of property. To the disinherited of the twentieth century, the landlord is no more than a particularly nasty kind of capitalist.

What resources do we include in land? Generally speaking, all the free gifts of nature which yield an income: agricultural and building land, mines, fisheries, etc. We do not include sunshine, rain, and other natural agents which assist in production, since these are not in anyone's ownership and control. They are resources over which we have no power of disposal, and which, therefore, we cannot economise.

The characteristic of land on which economists have laid most emphasis is the fixity of its supply. Coastal erosion and flooding may reduce somewhat the land surface of the globe, while slight additions may be made through exploration, discovery, or the building of dykes. But, broadly speaking, and with the possible exception of mineral deposits, no great change in the supply is likely to take place. If, however, we are thinking of fertile, cultivable, land this is much less true. Erosion, for instance, can lay waste millions of acres of arable land. This has happened on an enormous scale in the dustbowl area of the United States, where half a million acres of prairie land have been turned into desert. Continuous cropping, neglect, incompetence, or ignorance can all reduce the fertility of the surface soil. On the other hand, fertility can be increased by laying out capital on land improvement. Reclamation, drainage, manuring, irrigation—even afforestation—can make available a larger supply of cultivable land. Rich crops of cotton can be grown on land where sagebrush and cactus were once hard put to it to survive. It might seem, therefore, that to say that our land resources are fixed in supply is exaggerating a little. Fertility can be used up and it can be increased. There are no "original and indestructible powers of the soil," or if there are, it is impossible to disentangle them from the powers which capital expenditure on land improvement has created.

From this it is an easy step to the conclusion that there is no fundamental difference between land and capital, and that there is no need to classify them as separate factors of production. Land, it is sometimes argued, is simply a piece of property with a very long life, and of a sort not easily added to. There is no need to separate it too rigidly from other pieces of property by pretending that it can never wear out and never be added to. Its supply is not rigidly fixed, but it takes an unusually long time for an increase or decrease to take place.

All this is true if by land we mean cultivable land, or simply fertility. But when a farmer uses land (and still more when a builder does) he has the use not only of the powers of the soil (indestructible or not)

but also of a given area which enjoys an annuity of air, sunshine and rain and is in his exclusive occupation. Fertility can be changed by the farmer, but climate and situation cannot. We can improve site-values, but we cannot multiply the sites themselves. As for mineral deposits, it is quite plain that we can do nothing to increase the supply, although we can do a great deal to make the supply more accessible (e.g., by prospecting, sinking shafts, building railways, etc.).

Thus in land as we find it there are two elements which in practice can rarely be separated, but in theory always. One is variable and the product of human effort, and one is constant and the free gift of nature. It is the constant element—the advantages of climate, aspect, situation, etc., that come from the exclusive ownership and use of a site or of mineral deposits—which constitutes Land.

It is this theoretical distinction that has led economists to treat Land as a separate factor of production. But even if the distinction had never been drawn, it would almost have been necessary to invent some imaginary factor to take the place of Land—a factor which, like Land, would have been fixed in supply. In economics, we frequently have to distinguish between resources which are strictly limited in supply over the interval of time under discussion (e.g., machinery and plant), and resources which can readily be varied (e.g., semi-finished goods, labour, etc.). The first set of resources raises problems very similar to those raised by the use of land, and it is very useful to be able to apply the theory that has already been worked out in connection with land.

To take a simple example. An important corollary of the fact that land is fixed in supply is that the owners of land are in a position of monopoly. They have exclusive use of resources without which our most urgent wants cannot be met. If those wants become more urgent (for instance, because population increases and more people have to be fed), no new land resources can be created to supplement the old, and landowners are put in a position to hold society up to ransom by charging higher rents. They profit, not because they have performed some additional service to the community, but because they happen to own resources which have become scarcer relatively to the community's need of them. They will continue to profit because the scarcity of land cannot be relieved by setting men to work to make more of it.

All this is true, not just of land, but of anything that is fixed in supply. If, for instance, there is a sudden, but sustained, demand for steel from the shipbuilding or automobile or armament industries, it will be impossible to expand the capacity of the steel industry for some time. Until new furnaces and rolling-mills come into production, the owners of the existing steel plants are in the same position of advantage as landowners faced with a rising demand for land. They are able to raise their price to the higher level which the market will bear, and reap an extra profit without extra effort, But theirs is not a continuing advantage, for the scarcity on which they batten is one that can be modified. After a time, the high profits will encourage the

construction of new steel-making plant, which will enter into competition with the old and bring down prices and profits. In the long run, the supply of plant can always be increased: high profits, therefore, tend to be self-extinguishing. But even in the long run, the supply of land cannot be increased: high rents, therefore, may continue indefinitely.

A second characteristic of land is that it has no cost of production. No one paid a penny to have it created, and it costs no one a penny in actual outlay to allow it to be used. It is already in existence and awaiting employment. In this it differs from both labour and capital. Labour has to be reared—a troublesome and costly business—and does not offer itself for employment without a struggle; there are sacrifices to be borne by the parents of a worker in bringing him up and by the worker himself in seeking employment rather than yielding to the attractions of leisure. Similarly, capital has to be built up out of savings—which for most people means a sacrifice of present enjoyment. It is not, like land, indestructible, and can be used up and not replaced if the return to investment is not satisfactory. Thus while the provision of labour and capital is costly—in the sense that it puts us to some sacrifice either of leisure or of present enjoyment—there is no cost, no sacrifice of some desirable alternative, in providing land. The land is there for the owner to make the best of it, but labour and capital are not "there." How much of them exists depends largely on what they are paid, not on nature's bounty. We are not committed to saving some given sum, or working for so many weeks, or days, or hours at some fixed speed, irrespective of the efforts and sacrifices involved. We count the cost of working and saving. But the productive power of land is not of our making and puts us to no cost. If we choose to improve this productive power by sinking capital in land, then such man-made fertility has, of course, a cost of production. But natural fertility and the advantages of climate and situation possessed by any particular piece of land have no cost of production.

From this it follows that an increase in land values must represent an increase in the wealth of individual landowners without the performance of any equivalent service on their part, unless the rise is due to the sinking of capital in improvements. On the other hand, a fall in land values represents a windfall loss to landowners with no offsetting gain. Generally, with an increasing population, we expect to see land values moving upwards. In nineteenth century Britain, however, land values remained remarkably steady (taking agricultural land only).[1] This must be set down to the opening up of vast tracts of land in America and the Empire, and the consequent damping down of the pressure on British land. The land of the New World came into competition with the land of the Old, drawing off large numbers of emigrants, and sending back supplies of foodstuffs which,

[1] Rents moved up from £42 m. in 1815 to about £60 m. in the late 'seventies, and down again to £42 m. in 1914.

instead of becoming increasingly dear, became increasingly cheap. Nevertheless the fear remained, and before the last war was gaining ground, that the needs of an increasing population would eventually force up land values and deposit an " unearned increment " in the lap of British landlords. As for *urban* land values, they rose steadily as the towns grew in size, and there was never any doubt that the fortunate owners of sites in and around the towns were enjoying a substantial " unearned increment."

The fact that it costs nothing to supply land (as contrasted with the cost of labouring or saving) leads to a second important conclusion. For if it costs nothing to supply land, nothing is gained by making no use of it. Hence if the choice lies between letting land for an almost nominal rent and not letting it at all, it will pay to let it. The rent will be fixed by what the market will bear, not by cost, for the reason that there is no cost. It will always pay to obtain from land whatever income it will bring in, over and above the expense of farming it. Where land is not used for any purpose whatever, it is presumably because it is too inferior in fertility or in situation to yield a surplus above the cost of farming it.[1]

This reasoning can be extended to any productive agent where nothing is gained by failure to make use of it. A piece of machinery, for example, may depreciate no more rapidly when in use than when idle. If so, the real cost of using it (the sacrifice to which its owner is put) is nil, and it will pay to keep it in use so long as operating costs are covered. It is true that, unlike a piece of land, the machine cost something to produce, that the owner has hopes of recovering this " sunk " cost, and that, if he does not, he will be unlikely to replace the machine when it wears out. But his past hopes and present chances are two different things. The cost of producing the machine has little connection with the cost of using it. The first was incurred long ago, and in economics one of the most important principles is that " bygones are forever bygones." The second is the cost of efforts and sacrifices which people are induced to make *now ;* and the use of our machine involves no sacrifice whatever. The cost of using it is nil. We have here one more example of the analogy drawn above between land and durable capital.

A third characteristic of land is its heterogeneity. No two pieces of land are exactly alike in fertility or in situation. Some are highly fertile and situated near large urban markets, while some are on mountain tops and miles from anywhere. It is possible to arrange each piece of ground (or each deposit of minerals or each fishery) in descending order of value, ranging from those which it would pay to use in almost any circumstances to those which are unlikely ever to be used at all. We might have Manhattan Island at one end of the

[1] In the real world, where agricultural land depreciates when not in use (e.g., because it becomes overgrown with weeds) the reasoning given above is reinforced. It is doubly important to find a user for land if capital is lost as well as no rent earned when the land lies idle.

list and the South Pole at the other. A line can be drawn somewhere on the list between pieces of land which repay cultivation and those which do not; between pieces which are worth building on and pieces which are not; between deposits which can profitably be mined and deposits which can not; and between fisheries which it does and does not pay to exploit. This line is generally referred to as *the margin*. Land for which there is no remunerative use of any kind is called "sub-marginal"; land which it is just worth while to cultivate is said to be "on the margin of cultivation"; and land which yields a substantial surplus above farming costs is described as "intra-marginal." The margin, or line of division, is not, of course, fixed. It is pushed outwards, so as to include land that was formerly sub-marginal, if, for example, population increases and the demand for foodstuffs becomes more pressing. Or if, to take the opposite case, population falls, or people eat less food, or larger crops can be grown on the existing area under cultivation, the margin will press inwards and more land will pass into the sub-marginal, unused class. Similarly with mining. A fall in mining costs, for example, will bring pits which were previously uneconomic above the margin, while a rise in costs will force some pits over the margin into idleness, because mining operations have ceased to be worth while on the thinnest seams, or leanest ores, or deepest deposits, or in the most gaseous or flooded pits.

The margin of cultivation divides land which yields a surplus over costs of cultivation in any use (no matter what), from land which cannot be made to yield a surplus in any use whatever. A second kind of margin, the margin of transference, divides land which it pays to use for one purpose (e.g., dairy-farming) from land which it pays to use for another (e.g., cattle-rearing). We can range land used for dairying and cattle-rearing in order of what one may call "relative suitability" for dairying, at current prices of milk and store cattle and at current levels of farming costs. There will be some land (e.g., rich pasture or land near the towns) for which dairying can easily outbid cattle-rearing. There will also be some land (e.g., hill-grazings and land in rather inaccessible places) where the advantage clearly lies with cattle-rearing. Land intermediate between these types can be allotted with more difficulty, and may pass from one use to the other, crossing and re-crossing the margin of transference, as milk prices rise or fall relatively to cattle prices. Land near to the margin of transference may be relatively well-suited to both dairying and cattle-rearing (and pay a high rent), or relatively ill-suited (and pay a low rent). The main point is that the pull on both sides, whether strong or weak, is roughly equal. If the pull of one side is strengthened (e.g., by a rise in the price of milk) more land is transferred across the margin, and a new margin, at which the pull on both sides balances again, comes into existence.

If pieces of land are different from one another, so are workers and their tools. The third characteristic of land is one shared in by all the factors. The conception of the margin, which derives from this characteristic, can be applied, therefore, to labour and capital

as well as to land. If there are marginal tracts of sheep-land, there are also marginal shepherds and marginal sheep. There is no conception in modern economics of which such extensive use is made, or which is so fundamental to an understanding of the subject, as the conception of the margin. We shall meet with it in many forms from now on.

**The Law of Diminishing Returns.**

Reflection on the characteristics of land gave us, about 150 years ago, one of the most famous of economic laws—the law of diminishing returns:

*Successive applications of labour and capital to a given area of land must ultimately, other things remaining the same, yield a less than proportionate increase in produce.*

This "law" is simply a generalisation based on experience. If the law were not true, if by doubling his outlay on labour and capital a farmer could double his produce, every farmer could save nearly the whole of his rent by giving up all but a small piece of his land and concentrating all his labour and capital upon that piece. Instead of spending £10 on each of 50 acres, he could multiply his outlay fifty-fold on a single acre and still grow as much as before. Similarly there would be no point in having large herds of cattle if every addition to the feeding stuffs increased the weight of the cattle at a steady rate, without limit. A single bullock would supply the nation. But the bullock, in point of fact, responds with constant or increasing returns only if he has been half-starved, and with rapidly diminishing returns if he has not.

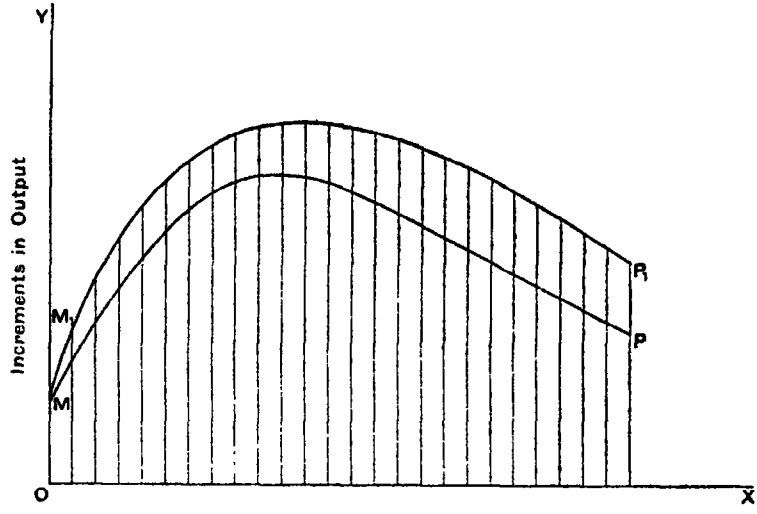
A farmer may, *for a time*, work under increasing returns (that is, additions to his outlay on labour and capital may yield a *more* than proportionate increase in produce). It may happen for example, that he spreads his work over so large an area that he would gain by concentrating his labour and capital on a smaller space. This may be the result of bad organisation; he may, for example, be sowing so lightly that his crops are smothered by weeds. But it may also in the short run be due to the inability of the farmer either to use less land or more capital. His land may, for example, be understocked, in the sense that if he could raise the necessary capital he could feed a larger herd of cattle and obtain a larger return per pound of capital invested. In the same circumstances it would pay the farmer to use less land (and so save rent) but the terms of his lease may make this (for the time being) impossible.

Other things, too, must remain the same. The return to labour and capital per acre of land may be raised and the law of diminishing returns suspended if some new discovery makes land more fertile, or if, because farmers are better trained or more skilled, they are able to put capital to a better use than their less efficient predecessors. Fertilisers can treble the weight of wheat crops: new brands of sugar can yield six times as much as the old: new feeding stuffs can fatten cattle to weights undreamt of 200 years ago. But once a discovery

has been made the law reasserts itself. At Rothamsted wheat was grown on five plots, one receiving mineral manure alone and each of the others various " doses " of nitrogen. One dose of nitrogen increased the yield by 10.3 bushels, a double dose by 19.4 bushels, three doses by 21.3 bushels and four doses by 21.8 bushels. The successive increments of nitrogen therefore conformed to the law of diminishing returns by yielding diminishing increments of product.

Many of these propositions can be more readily understood with the help of a diagram. In Diagram I, units of labour and capital applied to, let us say, an acre of land are plotted on the X axis and the increments in product yielded by each successive unit of labour and capital are plotted along the Y axis. The total crop is represented by

DIAGRAM I.




Outlay on Labour and Capital.

the sum of the increments, i.e., by the total area under the curve MP At first, there are likely to be increasing returns: if twice as much labour and capital is used, the crop will more than double. Later, there will be decreasing returns: the increments (represented by strips of constant width) will become smaller and smaller. If an improved type of fertiliser comes on the market, the productivity of the soil will be increased, and the curve of productivity will rise from the position MP to the position $M_1P_1$. If labour and capital become scarcer and costlier, the distance along the X axis which we will move for each £1 spent will shorten (i.e., the strips will be narrower) and the same crop will now cost more to grow. It will be observed that the diagram shows actual quantities of product and not the cash value of these yields.

The law of diminishing returns is not strictly applicable to mines. No doubt the application of more and more capital to mines results in a diminishing rate of yield. There is a continually increasing difficulty in finding fresh supplies of minerals, and more and more inaccessible stores have to be drawn upon. But the yield of mines, unlike the yield of land, is not a constantly recurring income. Land properly cultivated retains its fertility; the produce of the mine is part of the mine itself. Capital and labour can be used to economise land by cultivating more intensively. But when mines are worked at greater depths or in less rich seams, there is no question of *economising* minerals. The minerals are there and must either be mined or not mined. The distinction here between mines and agricultural land is parallel to the distinction between mining royalties and rent. Royalties are levied in proportion to tonnage mined; but rent is paid *annually*, and the farmer contracts to give back the land as rich as he found it.

The law *is* applicable to building land. The principles governing ground rents and farm rents are fundamentally the same. The more capital is applied to a vacant site, the more accommodation is available. But each successive storey, beyond a certain height, is less and less convenient, and more and more costly to build. We cannot build a city vertically like the Tower of Babel any more than we can grow all the world's wheat in a back garden. In the end, therefore, builders are forced to pay more rent for extra land and push out the limits of the city just as farmers, when they reach a point where more intensive cultivation ceases to pay, prefer to apply their capital to a larger acreage.

A tendency to decreasing returns is generally associated with a tendency to increasing cost. If equal applications of labour and capital yield diminishing increments of product, then each successive increment will be increasingly costly. The fifth pound's worth of fertiliser may add two bushels to the yield of an acre of wheat while the tenth pound's worth adds only one bushel. The cost of an extra bushel, therefore, will rise from 10/– to £1 as cultivation becomes more intensive. In other words, the greater the quantity of foodstuffs which we try to raise on a given area of land, the more costly will these foodstuffs be. It is this conclusion, so enormously important to a country with a rapidly increasing population, that has made the law of diminishing returns so celebrated.

For a clear understanding of the laws of diminishing returns and increasing cost, it is essential to keep two points in mind. First, the laws have nothing whatever to do with *profits*. They relate only to cost, or to returns in the sense of physical product. In discussing the laws, our attention is concentrated on the production of goods without bothering about the price which the goods will fetch. Secondly, the laws do not tell us what will in fact happen to costs if output increases over a period of years. The cost of manufacturing a commodity may be falling steadily because of inventions, discoveries, and so on. But these influences are excluded by the clause "other things remaining the same." The fact that the cost of growing wheat *is becoming* less, or *will be* less ten years from now, is irrelevant to the

question whether, under present circumstances and using existing methods of cultivation, wheat is being produced at increasing cost or not. If each additional bushel costs more to produce, other things remaining the same, then, whatever does happen or has happened in wheat-farming, the law of increasing cost is in operation.

### 3. Capital.

Of all the factors, this is the most confusing. Controversies about what capital is and does, what should be included in it, and what excluded, are endless. The niceties of controversy, however, have no place here. Our task is the simpler one of clearing away some of the grosser misconceptions, formulating some of the more obvious distinctions, and pronouncing, without excessive dogmatism, on points of major importance about which economists still differ.

When we speak of our capital we include three different things: our goods and chattels, our money, and our titles to wealth (stocks, shares, bonds, etc.). The first is concrete-capital, the second, money or loan-capital, and the third, debt-capital. Let us take these in turn.

**(a) Concrete-Capital or Property.**—Concrete-capital consists of a stock of assets possessing a money value. In the narrow sense of trade capital, it includes only assets in the hands of producers; in the broad sense, it becomes co-extensive with social wealth, and includes assets in the hands of consumers (dwelling-houses, motor-cars, etc.) or belonging to the community.

Trade capital consists of the fixed instruments of production (buildings, ships, machinery, etc.), goods in process, and stocks of finished goods. The last two items are classed together as working capital, while the first is called fixed capital. The line of division is easily drawn, since fixed capital is not used up in a single use, whereas working capital is.

Social capital is something rather broader. It includes not only trade capital, but also non-commercial assets that possess a money value. These are productive resources equally with trade capital, since they have been accumulated with a view to satisfying human wants. We cannot, for example, rule out house-property from capital; for it is just as indispensable to meeting some of our wants as are railway carriages or blast furnaces to meeting others. Nor can we draw a satisfactory line of division between houses let to tenants, and bringing in an income of profit to their owner, and houses retained by landlords for their own use and supplying them with shelter which would otherwise have to be bought.

The distinction between social capital and wealth is one of standpoint. Capital is an agent in *production*: it represents resources that can be used in the future. Wealth is a fund upon which we can draw in *consumption*: it represents stored up facilities for the satisfaction of future wants. But what are "resources" to the producer are "facilities" to the consumer. No one distinguishes as a private citizen between his capital and his wealth, except when speaking of his

stock-in-trade (i.e., of trade capital). For society as a whole, the same identity exists so long as we think consistently of wealth and capital in real terms as a stock of assets, and ignore debts altogether.

Arguments about concrete capital very often assume that the bulk of the capital of the country consists of machinery. When we turn to the calculations of the late Lord Stamp and other statisticians we soon find that this assumption is quite false.[1] We find, for example, that house-property alone forms over a quarter of the National Wealth, and investments in the British Empire and in foreign countries about a fifth. There is probably more capital sunk in the roads and railways of Great Britain than in all the machinery used in manufacture and mining. It is into buildings (including factory buildings, shops, etc.), into transport (roads, railways, ships, etc.), and into public utilities (gas, water, electricity, harbours and docks, etc.) that by far the largest proportion of our capital has gone. Even if we are thinking of current *additions* to capital out of savings, machinery rarely forms over a quarter of total new construction.

**(b) Money-Capital.**—So accustomed are we to include money in our capital that we do not always distinguish carefully between money and capital. The terms, in everyday speech, are used interchangeably. For instance, we speak of a millionaire "leaving a great deal of money." No doubt, by *our* standards, he did leave a fair amount of cash to his heirs. But his legacies probably consisted in the main of valuable property (ships, factories, rubber plantations, etc.) or of stocks and shares and other titles to wealth. Capital can always be given a money-valuation. But there is a difference between reckoning capital in terms of money, and using the terms "capital" (or "wealth") and "money" to mean the same thing.

When we seek to increase our capital, we do so first by adding to our stock of money. Normally it is money that we save and money that we borrow.[2] Often, therefore, we talk loosely of wanting more money when what we really want is more capital.[3] We want more money, not simply for ourselves, but, in our patriotic moods, for the whole country. We believe—quite mistakenly—that what is true for each of us must be true for all taken together, and that everyone would be better off if there were more money in the country. If there were more *capital* in the country, there would be a real gain to the community. But an increase in the stock of money would not, of itself, increase the stock of capital one iota. The countries which, after the war of 1914-18, were so successful in inflating their money-supply, did not seem satisfied that everyone was increasingly more prosperous. An increase in the supply of money—as David Hume argued two centuries ago—simply makes money less scarce and, therefore, less

[1] See above, page 17.

[2] We borrow *some* concrete capital (e.g., lawnmowers) and hire a great deal more (e.g., houses, motor-cars, etc.).

[3] Or, very often, a larger *income*.

valuable. As the increased supply circulates, prices are driven up and money loses purchasing power. To be at pains to increase the supply of money (at any rate in good times) is rather like forcing food on a thin man who is already well-fed. Instead of making him fatter, you simply ruin his digestion.

The fact is that *any* stock of money will do to let the business of the world be transacted if prices and incomes are in line with the existing supply. An increase in the total stock is of service to the community only in special circumstances, whereas an increase in a person's stock of money is of service to him at any time. Money is capital to a person because it represents liquid resources which he can turn into concrete capital whenever he chooses. Money is not capital from the social point of view because society cannot use it to increase its stock of concrete capital, any more than Robinson Crusoe could have done so. Unlike a person, society cannot dispose of its money because there is no one to whom the money could be disposed of. A single country, however, can dispose of its gold holdings (and sometimes of part of its currency) to other countries, and might fairly include this part of its stock of money in its National Wealth.

The confusion of money with capital and wealth might seem too crude to have played much part in history. But when we study the history of trade policy, we find that for centuries statesmen defended a policy of tariffs and trade restriction by arguments built on this confusion. They passed laws prohibiting the export of good English money under penalty of death. They took measures (such as tariffs on imports, and bounties on exports) to preserve a favourable balance of trade with foreign countries, and to force them to pay in treasure for the balance of their purchases. They believed, with Thomas Mun—an economist of the seventeenth century—that "the ordinary means to increase our wealth and treasure is by foreign trade, wherein we must ever observe this rule: to sell more to strangers yearly than we consume of theirs in value." This balance of trade doctrine chimed with the intense nationalism of the period—as it does still—and issued in beggar-my-neighbour policies which were designed to increase our stock of the precious metals and deprive other countries of theirs. These policies probably owed more to nationalist sentiment than to bad logic. But if the logic had been better—if there had been a clearer perception of the relationship between money and capital—the nationalism might have been more easily bridled. For it would have been seen that the money brought to this country by a favourable balance of trade simply drove up prices and raised the cost of living.[1]

We have seen that money-capital is not a factor of production in the sense that it forms part of our productive resources. And yet it is clear that money is constantly being converted into the concrete capital which does form part of our resources, and that concrete capital is always convertible and is constantly being converted into money. A manufacturer borrows money with which to build his factory (i.e.,

[1] For a fuller discussion of the balance of trade, see Chapter 27.

he converts money-capital into concrete-capital). A shopkeeper tries to sell his stock as fast as he can for money, with which he then buys more stock (i.e., he turns concrete-capital into money and back again). A great many decisions about capital, therefore (e.g., whether we will add to it, whether we will hold it in a liquid form, and so on), are to all appearances decisions about money. In fact, however, they are really decisions about how we will use our real resources. It is just as if money were a moveable fence round a large estate and we had to decide where the fence should stand. We should have to move the fence whenever we wanted to make the enclosure bulge at one point or retreat at another. Our decisions would appear to be decisions about the fence, but would really be decisions about the area and shape of the enclosure. Similarly with money-capital. Our decisions to save and invest money are really decisions about how much concrete capital there is to be, and what form it is to take. So long as we hold money, we can dictate whether resources will be used to meet our immediate wants (e.g., by spending the money on opera-seats) or whether they will be used in order to make an addition to our stock of capital assets (e.g., by investing the money in a new house). We have "free" capital at our disposal. Once invested, however, capital ceases to be "free." It is sunk in a real asset like a house, and can only be "recovered" (i.e., our money—or "free" capital with its power of disposal over liquid resources) as the asset wears out and the funds set aside annually to cover depreciation mount up to the sum originally invested.

To make all this clearer, let us see in more detail what is happening beneath the veil of money when we decide to save. First of all, we refrain from spending money on services (e.g., on tram fares, cinema tickets, etc.) or on goods for immediate consumption (food, clothing, etc.)—that is, on what are called "consumption goods." This means that we refrain from buying the services of those who make these goods. If we cut out buying chocolate, we deprive of a job the men, machines, and land which formerly supplied our want for chocolate. The first effect of saving, therefore, is to cause unemployment. Or, if as much chocolate as before is produced, stocks of unsold chocolate will accumulate, or will be disposed of only at unremunerative prices which impoverish the producer; neither of these things can go on indefinitely.

Presumably, however, our savings are borrowed and expended on capital extensions. The borrower draws into employment men, machines, and land, paying them out of our savings for the building of "capital goods" (factories, machinery, houses, etc.). Thus a boom in the capital goods industries accompanies and balances a slump in the consumption goods industries. Saving throws butchers and bakers and chocolate-makers out of work; investment creates employment for bricklayers, riveters, engineers, etc. The savers accumulate claims to the services of the factors of production; the borrowers exercise these claims. The transfer of money from savers to borrowers is a transfer of claims on productive resources. More-

over if the claims are not exercised pari passu with their accumulation, the savings have been wasted. The whole purpose of saving is to release productive resources from the task of meeting our day-to-day wants in order to undertake the more important task of enlarging our stock of capital equipment. To save is to spare productive power; and if that productive power is not utilised when the saving is done, it does not survive to be utilised later. Saving, as Professor D. H. Robertson has well put it, is the one thing that cannot be saved. There is no sense in sparing what no one proposes to use.

If, therefore, there is a hitch in the conversion of savings into concrete capital, or if productive resources cannot be easily switched from the consumption goods industries to the capital goods industries —and these are big " ifs," as we shall see when we come to discuss unemployment—saving has failed (in part at least) in its purpose. Instead of enriching the community by permitting works of enduring value to be set in hand, it will impoverish all those whose job it was to provide for the savers. The goods that are saved are not produced—there is no market for them ; and no goods are produced in their place.

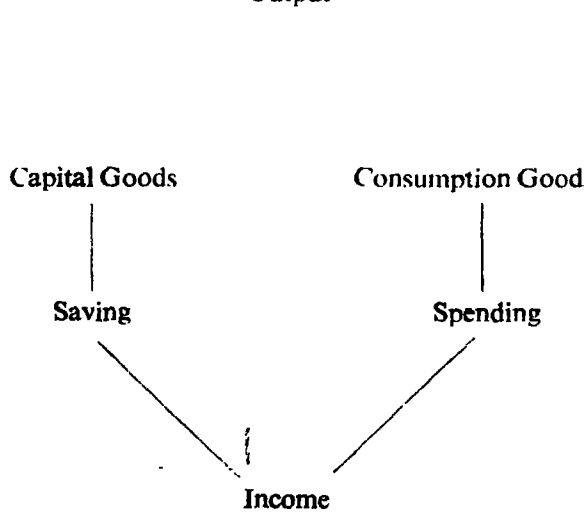
It is not just when savings are running to waste that the goods " saved " are not produced. As we have seen, what is saved is ultimately productive power, which is diverted from making the goods formerly bought by the savers to making capital goods for the borrowers. As savings increase, a change comes over the organisation of the labour force, so that more men are employed in the capital goods industries and fewer in the consumption goods industries. *Current* saving permits an addition to be made to concrete capital. Hence men work on new factories, new roads, new ships, new machines, etc. *Past* saving has accumulated a huge stock of concrete capital, which has constantly to be repaired and replaced. Hence new factories and roads and ships are needed because the old ones are wearing out; while an increasing number of men are engaged as engineers, plumbers, masons, roadmenders, etc., in maintaining the existing stock in working order. Production becomes more " round-about." Instead of setting about making goods directly, men are called upon more and more to make goods to make the goods, or even to make goods to make the goods which make the goods—and so on, much as in the House that Jack Built.

**(c) Debt-Capital.**—When money is lent, the burden of finding an outlet for savings in productive enterprise is transferred to the borrower. It is he who builds and owns the real asset that is added to our stock of capital. The lender retains only an expression of debt, a title to wealth, from which he expects to derive an income. Stocks and shares, for example, will be added to his capital, while the companies in which he has invested build new factories or instal more machinery If we are thinking of capital as a factor of production it is clear that we will include the machinery and the factories and exclude the stocks and shares. But if we are thinking of capital, as we often do, as anything that yields us an income, debt-capital will interest us just as

much as real capital. If it were true that capital always yielded an income because it was a productive agent, then it would be easy to reconcile the two points of view. The income yielded by debts would be exactly equal to the productivity of the corresponding assets, and it would make little difference whether we regarded capital as income-yielding or as income-creating. But, in fact, debts are often unchanged when the value of the assets changes, or when the assets (like the shells for which the National Debt paid) are blown to smithereens.

**The National Output and the National Income.**

The aggregate joint product of a country's resources of labour, land, and capital is known as its National Output. This is, for any given period (generally a year), the money value of the goods and services which flow into being during that period and which are customarily exchanged for money. The National Output, so defined, must be equal to the National Income, which is the aggregate joint earnings of the factors of production. The National Output is the flow of goods produced while the National Income is the flow of goods consumed. If the community neither saves nor dissaves, these two flows are bound to be equal to one another, for no community can consume more than it produces and no community will go on producing more goods than people are willing to consume. If saving is taking place, National Output and Income will still be equal to one another, but the flow of output and of expenditure will be split into two parts:




It is perhaps easier to see why the net value of output and income must be equal by thinking in terms of a single firm instead of a whole community. What a firm receives from sales every year must be

distributed in wages, or profits, or in paying for raw materials, or it must be saved and put to reserve. Now what the firm receives is greater than the *net* value of what it produces ; we have to deduct the cost of raw materials, for it is only the value *added* to the raw materials that is produced by the firm. On the other hand, what the firm pays out exceeds the income of its owners and employees ; what is paid for raw materials is not part of their income. Since, therefore, receipts and payments, which are equal to one another, exceed net output and net income by the same amount (viz., by the cost of raw materials), it follows that net output and net income must be equal to one another. If we add up the output of all firms we get the National Output, and if we add up their income we get the National Income. It will be true of the country, just as it is true of a firm, that what it distributes in income to the factors of production is equal to the value of what the factors co-operate to produce.

It follows from this that we cannot increase the National Income without simultaneously increasing the National Output. Before we can consume more there must be more to consume. The suggestion which is sometimes made that we should be better off if purchasing power were to be increased is true, therefore, only if the increase in purchasing power is accompanied by an increase in production, and this increase will be possible only if some of our productive resources are either not being used to the best advantage or are not being used at all (that is, are unemployed). An increase in the employment or in the efficiency of labour, for example, will certainly increase the National Income ; but an increase in wages may not. Higher wages will give wage-earners more purchasing power. But other people—for instance, those who pay wages—may have correspondingly less ; or, if they are able to pass on the rise in wage-costs to the consumer in the form of higher prices, the increased purchasing power of the community may be absorbed in paying higher prices for the same quantity of goods and services. *Real* incomes may be no greater than before. Nor is this just a theoretical possibility. The famous " Blum Experiment " in France, under which money-wages were increased by 30 per cent. in 1936–7, raised the cost of living and the money National Income in almost exactly equal proportions, so that the real National Income remained practically constant. All-round wage-cuts may have equally little effect ; they appear to be more successful in *re-distributing* income between different groups inside the community than in altering the community's aggregate real income.

The proposition that the National Income and the National Output must be equal to one another is sometimes denied. It is suggested that there is never enough purchasing power to allow all that might be produced to find a market. A baker, it is alleged, generates income when he pays wages to his men, but not when he buys coal or flour. The result is that there is a permanent deficiency of purchasing power. There is never enough " money " (meaning money-income) with which to pay for the bread which the baker is trying to sell. The result is the chronic unemployment with which everybody is familiar. If the

banks were to issue more " money " (meaning cash) to make good the deficiency in purchasing power, unemployment would soon disappear. It is the selfishness of the banks, who have a monopoly of " money " (meaning credit), that is to blame for all our woes. Now although there may be a grain of truth in the conclusion to this line of argument, the argument itself is entirely fallacious. The raw materials used by one man are the finished product of another. The purchase of flour, therefore, is just as successful in creating purchasing power as the payment of wages. The miller who sells flour to the baker, and the farmer who sells grain to the miller, draw income from the payments made by the baker just as much as do his employees. Whenever a service is rendered and paid for, purchasing power is created and this purchasing power must, in the aggregate, be equal to the value of the services rendered. There is no permanent deficiency of purchasing power. There may be a deficiency of *demand*—if, for example, people do not spend all they earn. But this raises a quite different question, which will be discussed later. Meanwhile we must be very careful to distinguish between wage-payments and purchasing power and between money and purchasing power. If we want to increase purchasing power (in the sense of money-incomes) we can, of course, increase money-wages or we can create more money. But we might equally well increase dividend payments, or house-rents, or interest charges. Purchasing power (the National Income) rises when the National Output rises, and taking measures to keep up purchasing power implies taking measures to keep up output and employment.

What it is that decides how much use we make of the factors of production, and what causes unemployment, we must leave over for later discussion. We must also leave aside for the present questions relating to the share of each factor in the total output. In the meantime we can turn to simpler and less controversial problems and see how, in a capitalist community, the production of goods and services is organised.

# PART II—THE STRUCTURE OF INDUSTRY

## CHAPTER 4

## SPECIALISATION: (1) DIVISION OF LABOUR

### The Advantages of Specialisation.

WE make our living by specialisation and exchange. We specialise: that is, we limit the range of our activities. And we exchange, enlarging the range of wants which we can satisfy. We give up variety in our work to make sure of variety when we come to spend the proceeds. We may have a hankering for self-sufficiency and life à la Robinson Crusoe; but we are generally thankful that the hankering is never put to the test of experience. It was all very well for Crusoe with his abundance of cocoanuts and goats. But in an island like ours, overrun with people, and poorly supplied with fruit and nuts, no Crusoe could hope to enjoy the comfort, the shelter, the diet, or the leisure, which are taken for granted by the average British worker. It is only a couple of centuries since a bad harvest in Scotland used to bring famine and starvation to thousands. In those days, when exchange still played a small part in people's lives, men were at work on Scottish farms from dawn until dusk, were fed on oatmeal and very little besides, and lived, very often, in hovels of turf that had to be rebuilt every few years, when the old turf was laid out as manure in the fields. To give up specialisation now would condemn us to a life almost as precarious and mean; and it would probably condemn to death from starvation millions of men to whom the crowded land of Britain could not afford a living.

The enormous increase in our efficiency in the last two centuries is obviously not due solely to greater specialisation. At least three other important influences have been at work—invention, capital accumulation, and colonisation. These influences have, however, been tangled up with the general trend towards specialisation. The steam-engine, for example, could never have led to the railway had there been no division of labour between engine-drivers, engine-makers, guards, porters, signalmen, etc. Capital for railways could never have been raised had there been no specialists in finance—promoters, bankers, issuing houses, etc.—to investigate, and to obtain support for, business propositions which might be made to yield a profit. The opening up of new countries and new sources of raw materials called for specialisation between the old countries and the new, manufactures and capital equipment going out from the old countries in exchange for foodstuffs and raw materials. Without specialisation, the *advantages* of invention, of capital accumulation, and of colonisation would have

been much less. So, too, one may add, would have been the *amount*. There would have been far less scope for invention, far less riches out of which to accumulate capital, and a far slower development of colonies.

Specialisation, then, was a pre-requisite of improvement. It was more: it was also a cause. It had an independent contribution to make to industrial efficiency—and a very important one too. This contribution was first analysed by Adam Smith. He pointed out that no two persons have exactly similar gifts. A professor may be a poor hand at cookery, and his cook be incapable of lecturing coherently. It will be an advantage if each concentrates upon what he can do best. Even if the professor is an excellent cook it will still be an advantage (since professors are more difficult to replace than cooks) if the professor refrains from taking up cookery.[1] The *special* gifts (and not just *any* natural gift) of each man are given the greatest possible scope when he specialises so as to devote the whole of his time to the occupation for which he is best suited. The labourer can make full use of strength and endurance, the skilled artisan of manual dexterity, and the company director of good judgment and organising ability. At the same time, those who lack the skill or physique necessary for some kinds of work are able to concentrate on other tasks in which their deficiencies are less of a handicap. Jobs in which a healthy or talented man would be wasted can be left to cripples and dunderheads. But specialisation is never perfect: there are plenty of square pegs in round holes. People may misjudge the aptitudes which they possess or which their job requires. Or they may be prevented by ignorance or bad organisation or lack of opportunity from finding work which makes full use of their natural bent. It is true that the force of competition will reduce the number of misplaced specialists by penalising them for their "error." People who mistake their vocation generally earn less than they might. But it would often be much simpler to offer a little good advice ("vocational guidance"), or free education, or more efficient labour management, rather than depend on the workings of competition to make people gravitate to suitable jobs.

Specialisation not only frees men to follow their natural bent. It *develops* skill in the specialist. Even if all men were alike, they would acquire, by constant practice, a knack of doing their specialised jobs. "The division of labour," as Adam Smith put it, "by reducing every man's business to some one simple operation, and by making this operation the sole employment of his life, necessarily increases very much the dexterity of the workman." Movements which at first are performed slowly and with difficulty are relegated, in William James'

[1] There may be special circumstances, however, which make it worth his while to give at least some of his time to cookery: e.g., in self-defence, because the only cooks whom he can find are a menace to his digestion; or to indulge a passion for cooking; or to profit from a rise in the social importance of cooks and a general detestation of professors. There is a parallel between these cases and protectionist arguments in favour of limiting *international* division of labour.

phrase, " to the effortless custody of automatism," and become rapid and precise. A trained typist will work much faster and with far less effort than an amateur who types only occasionally. Practice also familiarises a worker with the difficulties of his job, so that, for example, he can cope quickly with breakdowns and take steps to avoid them. Improvements, too, may occur to him. As he performs the same tasks day after day, he may see some method of simplifying the machinery or making it more nearly automatic.

Next there is a saving in time. In switching from one job to another, workers have to put away one set of tools and materials and assemble another. Very often they have to re-set their machinery—for instance, to spin a different count of yarn, or weave a different cloth, or roll a different section in a steel rolling mill. More time is wasted in warming to the new job. In the change-over, a man " commonly saunters a little. . . . When he first begins the new work he is seldom very keen and hearty; his mind, as they say, does not go to it, and for some time he rather trifles than applies to good purpose."[1] In unspecialised work, not only time but equipment is wasted. A man can only engage in a large number of tasks if he owns a correspondingly large number of tools, or if he runs the risk of being unable to borrow some tools when he wants them. When work is specialised, however, each worker is able to put his tools to constant use, and needs a much smaller supply of them.

Finally, specialisation makes possible a great extension in the use of machinery. Any manufacturing process can be split up into a number of distinct operations. " There are about ninety operations in making a pair of shoes, a hundred and fifty in making a man's coat, and thirteen in making a strip of steel into a spring leaf for an automobile spring."[2] Of these operations some are simple and regular, involving only the automatic repetition of identical movements. These can be taken over by machinery. Other operations require adjustment and control by a human agent. So long as these cannot be dispensed with, mechanisation can never be complete. In the making of electric lamps, for example, the wire may be drawn, the glass blown, and the lamps sealed mechanically, but the building, feeding, resetting, and repairing of the machinery, and the collection and assembly of the parts, require constant control by a worker whose movements are never *exactly* repetitive.

Now if every stage of manufacture from start to finish were in the hands of the same worker, it would be very difficult to disentangle the mechanical operations and turn them over to a machine. For the machinery necessary for a single stage in manufacture is very often much too expensive to be worth introducing unless it is kept in constant use, and it is likely to be capable of handling an output far greater than a single worker could cope with. If, however, each worker makes himself responsible for one or two operations only in the manufacture

[1] Adam Smith : " Wealth of Nations," Book 1, Chapter 1.

[2] S. H. Slichter : " Modern Economic Society," page 93.

of a product, each stage of manufacture can be kept in gear with the succeeding stages, and if there is a wide enough market for the finished product, no opportunity of mechanising any operation need be missed. If a machine at one stage can handle a thousand times as much as the machine at the next, we need simply use the two machines in the ratio of a thousand to one, and adjust the number of workers at each stage correspondingly.

Thus the introduction of expensive machinery is generally worth while only if men will give their whole time to feeding it and tending it, and give up working on other products, and even at other stages of manufacture. The use of machinery, in other words, dictates a comparatively narrow type of specialisation. From the simple division of labour where each man confines himself to the making of a single product from start to finish (specialisation by *product*), we pass to a more minute division of labour where the worker is responsible for one stage only in the process of manufacture (specialisation by *process*). The range of the worker's tasks is narrowed and standardised to suit the needs of the machine, while his share of the finished product becomes more and more difficult to identify. He may be tightening a succession of bolts with an electric wrench, watching for dents on tin cans as they pass before him on a moving belt, or assembling the same component part on one wireless set after another. There may be nothing that he can point to as his handiwork, or only some standard part, interchangeable with the product of thousands of other men.

### Machinery and Efficiency.

Although machinery goes hand in hand with specialisation, the two things are obviously quite distinct, and each makes a separate contribution to industrial efficiency. The special contribution made by machinery can be summed up as precision and power.

**Precision.**—The products of the machine, like its movements, are identical (or nearly so). Thus all the economics of standardisation are opened up whenever mechanical production becomes possible. In the days of James Watt and Adam Smith, " screw threads were in the main cut by hand by chipping and filling, every screw was different from every other, nuts and bolts had to be paired by specific fitting, and replacement in the ordinary sense was impossible."[1] Nowadays watch-screws weighing nine thousand to the ounce can be turned out at a vast rate and are indistinguishable under a microscope. The parts of a motor engine are accurate to within one-thirtieth the thickness of a human hair. Thus they can be made and tested in different departments (or in different factories), and can be assembled without waste to make engines whose reliability is known in advance. The standard parts are interchangeable, and as they can be kept in stock everywhere (or obtained quickly from the manufacturer) repair and

[1] " Factors in Industrial and Commercial Efficiency," page 273.

replacement are cheap and easy. Expensive journeys are saved, since orders can be given from samples. An enormous amount of uncertainty is taken out of business: the product of the machine inherits the precision of the machine and maintains a common standard of performance on which the purchaser can depend. Above all, machinery can raise the productivity of unskilled labour enormously by putting at its command a precision unattainable by even the most skilled of workers.

**Power.**—Secondly, machinery harnesses natural forces to produce power immensely greater than the muscle power of men or animals. In the Industrial Revolution it was water power that was harnessed. A water-wheel, driven by the current of a stream, provided the motive power for the whole factory. Later, steam power was introduced, at first to economise water by pumping it back after use into the reservoir above the wheel, and shortly afterwards to replace water-power altogether. This was done by means of the steam engine which generated power from coal and started a new revolution in manufacturing and transport. Later still, the power to drive machinery was supplied by electricity generated from water, coal, and oil. Hydro-electric power alone enormously reinforces human labour-power. At Boulder Dam, in the United States, each of the fifteen hydraulic turbines can generate 115,000 horse-power—the equivalent in muscular power of about a million men. At the Grand Coulee Dam on the Columbia River the capacity of the power installation is over 2,600,000 horse-power. The electricity generated at these dams drives machinery in factories all over the western states of America. In British factories, where about three-quarters of the power used is electric (mainly steam-generated), each worker has the assistance, on the average, of 3 horse-power.

The power harnessed by machinery can take over tasks that exhaust the worker and shorten his life. Before the invention of the steam plane, for example, a carpenter employed to smooth floor boards was often an old man at forty. He soon contracted heart disease and did not last "in his utmost vigour above eight years."[1] The work of the nineteenth century navvy was tiring in the extreme. But the operator of a mechanical scoop, moving a huge load of earth some thirty yards every minute, can do the work of hundreds of navvies without feeling exhausted at the end of the day. The machinery, moreover, is tireless. It represents power on tap. In many kinds of plant it can be left to work on, practically untended, provided material is supplied at certain intervals, and the product taken away when finished. A whole chain of operations, in many of which human muscles count for nothing, can be delivered over to machinery to be carried through with almost infallible regularity and exactitude.

Finally, the force supplied by machinery enables us to do what would otherwise be beyond our powers. We can build and operate

[1] "Wealth of Nations," Book I, Chapter VIII.

motors, steamships, and aeroplanes ; we can roll steel and crush rock ; we can use electric vacuum-cleaners, refrigerators, central-heating, air-conditioning, and so on : all of which would be impossible without the assistance of machinery.

**The Dangers of Specialisation.**

We have concentrated up till now on the advantages of specialisation and machinery. Industrial efficiency is improved, unskilled labour used to better purpose, and more goods or more leisure made available to society. Against these benefits, however, must be set many serious drawbacks. Some of these are due purely to specialisation, and others mainly to the introduction of machinery.

First of all, workers may specialise too narrowly. A change of occupation is often a form of relaxation. It frees the mind from boredom and strain by introducing variety into work. It is also a form of insurance against unemployment, since the worker can turn from one occupation to the other when it suits him. Above all, it broadens a man's character and outlook. Set a man a wide variety of tasks and he has a chance to develop initiative, judgment, and resource. Narrow his tasks, and you may narrow the man. The more specialisation limits the scope of a man's work, the greater the risk that it will stifle something in his character. Industry not only provides for men : it also *produces* men. And the great danger in excessive specialisation is that it tends to be far more successful in providing for men than in producing men of the right sort.

Secondly, from the social point of view, specialisation loosens the ties by which the community is bound together. Specialists have their peculiar background. They live in a world of their own, cut off from other specialists by their training, their habits, and above all by their interests. There is thus a constant danger of sectionalism : the specialists entrench themselves in organised groups in defence of their interests, and often in disregard of the interests of the community. Against this sectionalism, unity of culture is the only prophylactic. But in a society of specialists, where men's ideas are often unintelligible to one another, their tastes incongruent, and their moral standards uncertain, such unity is hard to achieve. For a common culture in which all can share, there must be a common framework of reference and common pre-occupations. But when specialisation is pushed too far, the common framework tends to be destroyed, and the common pre-occupations to be overlaid by other, more specialised, interests.

Thirdly, those who specialise in making a single product or in doing a single job run great risks. They depend on their fellows to buy their goods and services and to sell them in return the products which they lack. They hope to be able to win a livelihood by exchanging the goods which they can produce and do not want for the goods which they do want and which others can produce. If exchanges can be made quickly and readily, well and good. But exchange is often costly, or slow and unreliable. The parties may live at great distances

from one another, or much time may be taken up in striking a bargain, or there may be no regular dealings, so that it is not easy to find a buyer or a seller at short notice.

These risks and difficulties are multiplied enormously when production is undertaken in advance of demand, and when the circle of exchange widens to include specialists in every country in the world. When the village tailor makes shirts to order for his neighbours he runs little risk. But when a number of textile workers, most of them strangers to one another, co-operate to make a shirt in Lancashire in the faith that someone in Australia will want a shirt of just that size, pattern, and colour, and that his path and the path of the shirt will cross when he is in the right frame of mind to buy the shirt, the risks of the textile workers are obvious. They are not, perhaps, so enormous as they seem, for men are creatures of habit, and their wants do not change so rapidly that they cannot be predicted with fair confidence a few months ahead—even from a distance of several thousand miles. Nor are the risks so great as they would be in the absence of rapid transport (by steamship) and modern methods of communication (by cable and wireless). The steamship cuts down the time which it takes for the shirts to reach Australia, while cable and wireless services put Lancashire and Australia in close touch with one another, so that news is available in one country shortly after it is known in the other. This saving in time reduces the period over which forecasts must be made, and so makes it possible to foresee market requirements with greater accuracy and more certainty. But when all allowance is made, risks remain enormous. For the worker, no less than for the capitalist, there is no assured market. His specialised gifts are not certain of constant employment; his specialised job is not certain to yield him a steady wage; his chances of turning his hand to some other job are not certain to remain good. He is at the mercy of a future which he cannot foresee.

Even if a worker is consistently reasonable in his judgment of the risks which he runs in specialising, the unforeseen may occur and all the risks of specialisation will come home to roost with him in poor wages and unemployment. But he is not likely to judge reasonably. He will be tempted to over-specialise by ignorance of the risks involved, or by over-confidence in his ability to retain a well-paid, specialised job, or because he knows that, when unemployed, he can still draw unemployment benefit. Nor will he have the opportunity, if he decides that specialisation is risky, of showing a preference for a wider range of tasks than he combines in his daily work. He has, at best, a limited choice of jobs, all of them, with few exceptions, highly specialised, and it will be hard for him to find the right niche in an industrial system catering so exclusively for specialists. It is not on the preferences of workers and on their judgments of future prospects that the degree to which they specialise mainly depends, but on the decisions of their employers and on force of circumstances.

The work of specialists, then, is co-ordinated through exchange. Exchange brings with it risks which ought to limit specialisation, but

which, through ignorance or over-confidence on the part of the specialists, or because of their powerlessness as wage-earners, do not in fact limit specialisation as much as they should. Men are taught a single trade when they could, with advantage, learn two. They are kept constantly to a single department when they could, with advantage, be moved from one department to another. And they are put to a few tasks, when they could, with advantage, undertake more varied work. But the advantage, in each case, would be mainly to the worker and to society, not to the employer.

Co-ordination takes place not only through exchange, but also within each firm. Thus a specialist, when he is not his own master, is at the mercy, not only of market forces, but also of the whims of his employer. He has to work to orders under a discipline imposed from above. Since the introduction of machinery, this discipline has become far more rigid and irksome.[1] Each worker, whatever his temperament, must adjust his speed of working to the pace set for him by the machinery or by his mates. If he is a slow worker, the strain of speeding-up will tell on his nerves; if he is a fast worker, he will be annoyed and bored because he is constantly being held up and finds it impossible to work steadily.

**Machinery and Working Conditions.**

To the drawbacks of specialisation we must add the drawbacks of machinery. We have to put up with working conditions that are noisy, dangerous, and unhealthy. We have to allow the pollution of the atmosphere with smoke and fumes from lorries and engines, factory chimneys and power stations. We have to crowd workers together in the vicinity of factories, surrounded by din, dirt, and dinginess. Above all, machinery very often makes work highly monotonous or drives men out of work altogether.

**Monotony.**

We must be careful, however, not to exaggerate. It is commonly believed, for example, that the agricultural worker's job is far less monotonous than the factory-worker's. This is an attractive belief—particularly to the intellectual, with his horror of machine-tending and

[1] "The Industrial Revolution added discipline, and the discipline of a power driven by a competition that seemed as inhuman as the machines that thundered in factory and shed. The workman was summoned by the factory bell; his daily life was arranged by factory hours; he worked under an overseer imposing a method and precision for which the overseer had in turn to answer to some higher authority; if he broke one of the long series of minute regulations he was fined, and behind all this scheme of supervision and control there loomed the great impersonal system. . . . No economist of the day . . . ever allowed for the strain and violence that a man suffered in his feelings when he passed from a life in which he could smoke or eat, or dig or sleep as he pleased, to one in which somebody turned the key on him, and for fourteen hours he had not even the right to whistle." J. L. and B. Hammond, "The Town Labourer," pages 31-3.

his love of country life. But it is a belief, which, if we may take both sets of workers at their word, is almost certainly mistaken. There are few activities so monotonous as ploughing or milking[1]; on the other hand, factory work, with some striking exceptions, can be comparatively pleasant, especially when it means company and conversation. The operations of the cotton-weaver are often contrasted unfavourably with those of the old hand-loom weaver. But they are less laborious, less monotonous, call for more judgment, and are immensely more productive than his.[2] "I loved the mills," said an unemployed weaver to Mr. Oakeshott during an investigation at Blackburn. "I loved the company and the people and everything about them. The mill was home to me. I'd do anything in the world if I could get back to them."[3] Thousands of other workers, less outspoken, feel much the same of jobs which, to the outsider, seem dull and monotonous.

**Unemployment.**

Machinery creates unemployment; but it also creates employment. Labour is required for the building and installation of machinery, and to keep it in operation and repair. Presumably, too, the cost of producing the machine-made goods is reduced, and with it the price. Thus consumers are able to buy more of the machine-made goods, and more men are needed to produce them. Or if consumers prefer to buy some other kind of goods with the purchasing power which is released when prices fall, their expenditure will create openings for employment elsewhere. It will only be if the men whose jobs are taken over by machinery cannot, or will not, move to some other occupation that they will be permanently displaced.

This reasoning is confirmed by observation. Machinery has been introduced on an enormous scale in the last century and a half but the number of men in employment has increased five-fold. The industries which have the lowest unemployment percentages are often those, like motor-car production, radio manufacture, etc., which make the greatest use of machinery, while most of the industries in which unemployment is concentrated—for instance, shipbuilding and the

[1] cf. Disraeli: "There is no nation in the world that leads so monotonous a life as the Irish, because their only occupation is the cultivation of the soil before them. These men are discontented because they are not amused."

[2] cf. Marshall: "Principles of Economics," page 263. We should remember also that "many of those who perform the more monotonous parts of manufacturing work are as a rule not skilled workers who have come down to it but unskilled workers who have risen to it. A great number of those who work in the Lancashire cotton-mills have come there from poverty-stricken districts of Ireland, while others are the descendants of paupers and people of weak physique, who were sent there in large numbers early in the last century from the most miserable conditions of life in the poorest agricultural districts, where the labourers were fed and housed almost worse than the animals whom they tended." (*ibid.*, page 263 n.).

[3] "Men Without Work." A Report made to the Pilgrim Trust, page 151.

textile trades—are depressed mainly because of a deficiency of demand, not because they have been adopting more mechanical methods of production. As a rule, the workers who are displaced by machinery are absorbed within the firm which employs them. Of those who are not absorbed, a large proportion find work again within a short time. The unsuccessful remainder, who, in time of rapid technical change, may be fairly numerous, are left in chronic unemployment. But such "technological" unemployment, as we shall see later, is rarely more than a small fraction of total unemployment.

It is easy to idealise the past and pick on machinery as the cause of an imaginary fall from grace. But if we do justice to the immense improvement which machinery has made possible in health, education, and security, and in the working and living conditions of the mass of the people, we will be in no doubt that, on balance, machinery has made a great contribution to human welfare.

## CHAPTER 5

## SPECIALISATION: (2) LOCALISATION OF INDUSTRY

DISTRICTS, like persons, specialise. There is a geographical division of labour which leads each city, county, and country to abandon some lines of production and extend others beyond local requirements. Each area restricts the range of its products, because such specialisation enables it to obtain other products more cheaply by importing them. Through trade with its neighbours, the area is able to export what it has in comparative abundance in return for what it has in comparative scarcity.

Geographical specialisation can be looked at from two different angles. We may ask: Why does a given area carry on some industries and not others? Or we may ask: Why is a given industry carried on in some areas and not in others? Economists have built up the theory of international trade—to which we will come by and by—in answer to the first question, and the theory of the location of industry in answer to the second. In the study of international trade, our attention is generally focussed on the struggle which goes on between industries *inside* a country (or region) for the use of the scarce resources available there. We see how some industries survive while others are crowded out, and how the volume of industry expands if labour and capital are attracted to the country, and contracts if they flow abroad. In the study of localisation, we start from the competition *between* regions for a given industry, or for industry as a whole, and discuss why one region can pull harder than another. These two lines of thought do not conflict with, but supplement one another.

Suppose, for example, that we ask why some industry—say shipbuilding—has become localised on the Clyde. We have really two questions to answer, not one. First, why the Clyde rather than some other *place*—the Thames or the Severn ? Second, why shipbuilding rather than some other *industry*—steel or textiles ? We have to show, not only that there is a pull on shipbuilding to the Clyde, but that the pull is *relatively* greater than the pull on other industries.

That the Clyde has natural advantages in shipbuilding is obvious enough. The ship-yards are within a few miles of plate-mills and coal mines, and can obtain boilers, pump-feeding plant, deck and steering gear, fans and ventilating equipment, scientific instruments, and almost everything that is necessary in the fitting out of ships, from firms that have grown up with the industry along the banks of the Clyde. Skilled labour in abundance is available both in shipbuilding itself and in the subsidiary industries. Close contact can be maintained with the shipping companies which have their headquarters in Glasgow, so that the ultimate owners of many of the ships are able to exercise supervision over the building of them. Thus transport, labour, and administrative costs can all be kept low. It was these low costs which brought the industry to the Clyde and which have kept it there.

But if the Clyde is suited to shipbuilding, it is suited to other industries as well. Shipbuilding, after all, is a parvenu among Clyde industries. A century ago, the textile industries were by far the most important in the Clyde valley ; forty years later, the metal industries forged ahead ; and forty years later still, shipbuilding was in the lead. This transformation was not due to any change in the natural advantages of the district. The Clyde was, and remained, as well suited to cotton manufacture as Lancashire. There was an excellent port, an abundance of coal for the generation of power, an atmosphere of unquestionable humidity, and plenty of pure water for bleaching. The district was also well suited to the manufacture of steel and steel products. The coal measures lay next rich deposits of ore, and when these were worked out, pig iron and scrap could be imported cheaply by sea. The textile industries had trained up in the manufacture and repair of milling machinery a supply of skilled engineers on whose services the rising metal industries could draw. Finally, the nearness of Glasgow meant not only a local market for ship-plates, but also low freights on exports of steel to foreign countries. Yet both of these industries, without by any means disappearing, lost ground to shipbuilding. The Clyde ceased to rival Lancashire as a textile centre, and offered little or no competition to South Wales and the North-East Coast in the manufacture of structural steel, tinplates, and other steel products.

Thus it is not enough to explain the triumph of one industry over others, in a district like the Clyde, in terms of some natural advantage. A district may be well suited to a wide variety of trades, and yet concentrate on a few, just as a person with great natural gifts may develop only those in which he is most outstanding. Specialisation is not on the basis of absolute, but of *comparative* (i.e., relative) advan-

tage. The Clyde came to have a comparative advantage in shipbuilding when iron (and later, steel) displaced timber in the construction of the hull. This comparative advantage enabled the industry to attract labour from its rivals by offering wages higher than they could afford for comparable work: to attract capital away from them by offering higher profits: and to outbid them for the command of sites along the Clyde. Other areas, in which shipbuilding was on the decline, were at a comparative disadvantage in shipbuilding. The Thames, for instance, was well enough suited to the shipbuilding industry, but could not afford to pay the high wages which labour was earning in other industries in the London area.[1] The high costs which drove shipbuilding from the Thames were high because other industries were prepared to bid up the price of labour, land, and capital. The low costs which brought shipbuilding to the Clyde were low because other industries, less suited to the district, could not afford to bid up the price of labour, land, and capital. In each area, those industries which were at a comparative disadvantage were either squeezed out or forced to contract; they made way for industries which, by reason of their comparative advantage, could pay the highest prices for productive resources.

So far, we have been discussing the location of one industry—shipbuilding. We can now generalise the forces tending to bring an industry to a given area, and can set them out in a detailed list.

### 1. Natural Advantages.

First, there are the natural advantages of the area. These include its amenities—its possession of suitable sites and a suitable climate. Thus the flat, sunny country of East Anglia is far better suited than the wet glens of the Highlands to wheat-growing. Colour-printing is easier in a dry, and cotton-spinning in a damp, climate. And so on. The natural advantages of an area include also its access to sources of power, to markets, and to raw materials. Let us take these in turn.

**(a) Power.**—Before the eighteenth century, power-driven machinery was rarely used, and cheap power was not an important localising force. With the introduction of the water-wheel, factories—and especially cotton-mills—came to be built upon the banks of rivers with a good natural fall. Such rivers were more plentiful in the hilly parts of the country (i.e., in the north and west), and industries using water-power, therefore, tended to become localised there. A great impetus was given to this movement when steam came into use as the chief source of power, for coal, from which steam was generated, was located in the districts in the north and west to which the textile industries were already gravitating. Of course, coal might have been brought by rail to the older centres of industry, but as it was expensive

[1] Wages in the Thames yards in 1869 were more than 20 per cent. above wages on the Clyde. (J. Glover: "The Decline of Shipbuilding on the Thames," Journal of the Royal Statistical Society, 1869, page 272).

to transport, and none of its weight passed into the finished product, manufacturing costs were lower in the mining areas. As coal was used also as a fuel, as a reducing agent (in the metal industries), and as a raw material (in the chemical industries), the pull of the coal-belts was tremendous, and a whole network of industries grew up, attracting labour and capital from all over the country. There are signs that this process is now being reversed. The development of new sources of power—particularly oil and electricity—has greatly weakened the dependence of industry on coal, and is bringing about a redistribution of industry in favour of the large urban market (e.g., London). and at the expense of the mining districts (e.g., South Wales). The new factories which have sprung up round London in the last twenty years make use of electrical rather than steam power, and can buy electricity at least as cheaply as factories situated near the coal mines. At the same time, even when steam power and coal continue to be used in manufacture, the use of coal has been economised by great advances in furnace technique (e.g., in steel-making, in the generation of electricity, etc.), and the pull of raw materials and of the market has been strengthened relatively to the pull of coal. This has been illustrated very strikingly by the migration of the steel industry to the ore deposits away from the coalfields where it has traditionally been located. The old-established areas of South Wales, Scotland, and the North-East Coast have been losing ground steadily, while the Lincolnshire ores have attracted to Scunthorpe, Corby, and Redbourn steel plants engaged in the manufacture of sheets (Lysaght's), tubes (Stewart and Lloyd's), steel billets (Richard Thomas'), steel forgings and castings (Firth and Brown's), and other steel products. The migration of the industry is not due *solely* to changes in technique favouring location at the ore-fields. The expanding markets for steel products have been in the Midlands and the South (e.g., in the motor and building industries), while the contracting markets (either abroad, or in shipbuilding) have been those upon which the older areas are most dependent. It should be added, too, that the original location of the steel industry was due in part to the presence, in close conjunction to the coal seams, of rich ores which have since been worked out. In the United States, the parallel migration of the steel industry down to the Great Lakes, and away from the Pittsburgh area, has a similar origin. The industry is moving towards the ores of Michigan and Minnesota, and is following the market for steel products westwards.

**(b) Raw Materials and (c) Markets.**—Nearness to markets and to sources of raw materials very often pull in opposite directions, so that the advantages of one have to be weighed against the advantages of the other. Where there is great loss of weight in manufacture, or where the raw material is very bulky and the finished product compact and easily transported, production will be carried on near the source of supply of the raw material. Thus, the crushing and concentration of ores is normally undertaken at the minehead, even when smelting or refining is carried on elsewhere. The pulping of wood, the pressing

of grapes and apples, and the distillation of oil from coal are examples of the same tendency. If, however, the finished product is bulky, and loss of weight negligible, it will be the large markets which exert the biggest pull. The making of bricks, bread, bottles, and other bulky goods is generally carried on all over the country near each important market.[1] Tin-plate manufacture is more highly localised than any other industry; but the making of tins, cans, and household utensils from tinplate is widely diffused, because it is much cheaper to transport a box of plates than a pile of tins of various shapes and sizes. Similarly, the assembly of motor-cars parts is carried on in many countries in which the manufacture of these parts is uneconomical. The market has succeeded in attracting the later stages of manufacture because the final product is costlier to transport than the parts.

Where several raw materials are combined in a single product, and their combined pull is strong (because of loss of weight or bulk) in relation to the pull of the market, the final location of the industry will represent such a compromise as keeps transport costs at a minimum.[2] The importance of each constituent in relation to others will depend upon what weight of it is used per unit of the final product.[3] If it is a weight-losing material, its pull, relatively to that of *other materials*, will not be any the greater; but of course the pull of the raw materials vis-à-vis the *market* will be strengthened. The steel industry has moved away from the coalfields, not because coal loses less weight in the smelting of iron ore, but because less coal has to be moved per ton of ore.

When the raw materials used include some that are available everywhere (e.g., water, sand, clay, etc.), industry is attracted towards the market. For if these materials are incorporated in the final product at the point of sale instead of elsewhere, the cost of transporting them is saved. Brewing, brick-making, and, above all, the manufacture of aerated waters, illustrate the working of this tendency.

When several markets have to be supplied, the location of the industry will again represent a compromise between conflicting pulls. A large market like London, with enormous spending power per square mile of area (high "market density"), will pull a great deal harder than a country town or rural district. It will do so for two reasons: first, because production near a large market saves the cost of transport on goods sold locally, and these form a high proportion of output; and second, because large-scale production at a low cost is easier when market density is high. We have seen already that the pull of London and the South-East has been growing since the war because of the rise of new forms of power less highly localised than coal, and of new products less dependent upon localised materials. We have now to add that the increased importance of large-scale production and the

[1] Nevertheless bread-making is being increasingly centralised, and bricks are carried from as far north as Lairg in Sutherland to the London market.

[2] Other things (e.g., labour costs) being equal.

[3] And also, of course, on its situation.

greater ease and comparative cheapness of transport have tended to concentrate industry near the chief markets, and especially London, since the outlying markets can be supplied more cheaply from a single point of manufacture.

### 2. Acquired Advantages.

The natural advantages of an area in any given industry are generally reinforced by acquired advantages. Transport facilities are provided or improved. Roads, railways, docks and harbours are built. Commercial services of all kinds—banking, warehousing, accounting, and so on—become available. Labour becomes familiar with the technique of the industry, and a high standard of workmanship is developed. Manufacturers are able to discuss problems of mutual interest, and in this way, or because of the appearance of trade and technical journals, improvements and inventions are stimulated. The district acquires a prestige which binds customers to it. An organised market for the product or for materials comes into existence. Subsidiary industries spring up in the neighbourhood, supplying accessories and parts, or making use of by-products.

(a) **Vertical Disintegration.**—At this stage a change comes over the type of specialisation practised: as the industry obtains a firmer footing and grows in size, the process of manufacture often begins to "disintegrate," the separate stages as they split off being taken over by a specialist firm or industry. The supply of machinery and material, the utilisation of by-products, the organisation of traffic and marketing may be undertaken by such specialist firms. Cotton firms, for example, no longer make their own machinery but buy it from engineering firms. Motor manufacturers buy in radiators, carburettors, electric equipment, chassis-frames, and other parts from specialists who can produce very cheaply by supplying several firms at once. Shipbuilding yards assemble the products of dozens of industries. Similarly in building it is usual in large contracts for the steelwork, masonry, plumbing, air-conditioning, furnishing, and so on, to be contracted out to firms specialising in a single branch of the industry.

(b) **Lateral Disintegration.**—In a localised industry, "disintegration" by *product* is also frequent. That is, neighbouring towns (or firms) specialise in supplying different markets, or use similar materials to make different products. In the West Riding of Yorkshire, for example, Bradford specialises in worsteds, Dewsbury in "heavy" woollen cloths, Huddersfield in fine woollen cloths, and the Colne Valley in tweeds. In Lancashire, the cotton industry is even more highly specialised, the south concentrating on spinning and the north and east on weaving, while the towns inside these areas specialise within narrow limits in the counts of yarn which they spin or the fabrics which they weave.

There is no *necessary* connection between the industrial disintegration which we have been discussing and the concentration of an

industry in a given area. The first involves further specialisation on the part of firms and industries, the second, further specialisation on the part of geographical areas. In the world as we find it, however, both types of specialisation very often occur together; and it is easy to see that, so long as the firms in a given industry are isolated from one another instead of being localised in a single district, they will be forced to do many things for themselves rather inefficiently, and will have to do without services which a specialist could supply to a localised industry. *One* cause of disintegration (but not the only one) is clearly localisation. The more firms of the same type get together, the more types of firm do we get.

Acquired advantages not only strengthen the pull on an industry to a given area, but help to keep the industry there when the natural advantages of the area have ceased to exist. There is no obvious reason why lace-making should be carried on in Nottingham or boot-making in Leicester except that both towns have acquired a long experience in their staple trades. A group of interdependent industries is even harder to unscramble; for even when one of the group might profitably start again elsewhere it may be so knit together with surrounding industries that a change in location would be practically impossible. Changes in location, therefore, generally come about less through deliberate transference than through the rise of firms in a new area and the decline of firms in the old.

### 3. Relative Advantages.

Thirdly, the localisation of industry depends upon the relative advantages of each area: that is, upon the competition for labour, land, and capital within the area.[1] An industry will locate itself more readily in a district if other industries offer it little competition for productive resources. Thus, industries employing female labour may be attracted to the coalfields or to engineering towns, where there are few jobs open to women. A shortage of labour in one area may induce firms to set up branch factories in others. Industry will tend to move in search of cheap or subservient labour.

In the same way, districts with an abundance of capital and enterprise will attract new industries, especially if there is no counteracting locational pull to other districts. This abundance was common to all parts of the country so long as people were in the habit of investing in local industries and manufacturers could sell everywhere in a widening market. But the growth of investment through the Stock Exchange, the banks, insurance companies, and other investment intermediaries has concentrated a large proportion of new savings in London, while London manufacturers have been able to feed their enterprise on a large and expanding market for their goods. On the other hand, it has been difficult to invest or be enterprising in the Depressed Areas. Moreover, the preferences of business men (and still more of their wives) are rarely for life in a Depressed Area, and very often for life

[1] Compare the discussion at pages 53-4.

in London. These preferences, especially in the newer industries, are less and less over-ruled by other locational factors; they have aggravated the scarcity of enterprise and capital in the North and West, and have contributed to the drift of industry to the London area.[1]

Finally, the location chosen within any given district depends upon the competition offered by other industries for sites. Inside each town, for example, some industries will be willing to pay high ground rents so as to be near to their workers, or to docks, warehouses, and railway stations. Other industries, which have more difficulty in economising ground space and need ample room for expansion, may prefer to move to the outskirts where ground rents will be much lower. The clothing and milling trades are examples of the first tendency, motor-car and wireless manufacture of the second. The rough pattern into which the industries of a district sort themselves is in constant change as one industry yields space to others. One firm may vacate a site because the ground rent has become higher than it can afford: another firm may take over the site in order to extend its premises and retain a central location. There is a steady centrifugal pressure by which commerce squeezes industry from the inner parts of a town. The older factories near the centre are forced to re-locate themselves as hotels, offices, warehouses and shops creep outwards. The factories near the city boundaries may have to move farther out as the suburbs become built up.

### 4. Cumulative Advantages.

Once a great locational change is in progress and industry is on the move, the change will gain momentum from sympathetic movements of labour and capital. There will be not only a re-shuffling of industries inside each area, but also a series of expansions and contractions of whole areas, in the sense that more labour and capital will be put at the disposal of industries in expanding areas, and withdrawn from industries in contracting areas. This movement of resources will give the expanding areas a cumulative advantage, for it will both keep down the costs of firms trying to establish themselves there, and at the same time bring them a new local market when the immigrant workers begin to spend their wages. The contracting areas on the other hand will suffer from a cumulative disadvantage for they lose a customer in every migrant, and every customer lost may mean still another migrant.

[1] "Mr. Ford started to manufacture motor cars in Detroit because it was his home town. Sir William Morris (now Lord Nuffield) chose Cowley because the school in which his father was educated happened to be for sale. Neither of these excellent motives can be regarded as promising certain success to those who imitate them. . . . Mr. Ford and Sir William Morris either by acumen or by chance discovered places of production with great natural advantages." E. A. G. Robinson: "Structure of Competitive Industry," page 152.

**The Degree of Localisation.**—The more these advantages, whether natural, acquired, relative, or cumulative, are monopolised in any industry by a small number of areas, the more highly localised will the industry be. At one extreme, we have the mining of gilsonite (which takes place only in Utah) or nickel (which comes almost exclusively from Sudbury, Ontario); at the other, services of all kinds (of shopkeepers, policemen, postmen, lawyers, 'bus drivers, etc.) which, being in demand everywhere, and impossible to store, must be rendered everywhere.

**The Degree of Diversification.**—The more outstanding are the advantages of a single industry, or small group of industries, over other industries in any area, the narrower will be the range of industries in that area. In other words, industry will be highly diversified where the available resources, either in raw materials or in skill and experience, are varied, or when varied tastes and a wide market can be supplied; and it will be highly specialised when only a limited range of resources can be exploited, and when tastes are standardised and narrow. Chile, for example, cannot readily turn from guano to other lines of production, and the American prairies have little alternative to wheat-growing. But in Britain industry is amazingly diversified, not only because of our own varied resources—our coal and iron, our vast capital, long experience, and abundant skill—but also because our manufacturing centres, being within a short distance of the big ports, have easy access to the materials and markets of every other seaboard.

**The Depressed Areas.**—When the degree of diversification in an area is low, it risks all the disadvantages of excessive specialisation. If the main industry is depressed, the depression is communicated to the industries supplying local needs. There may be no other major industry to fall back on, and the difficulties of making a switch will then be enormously increased, since it is far harder to build up new industries than to build on old ones. Continued depression in the main industry will bring into existence a problem of long-term unemployment, especially amongst the older workers who can neither pick up a new technique nor move far afield in search of jobs. The younger workers may find employment in other areas, but their removal will unbalance industry more than ever. The social life of the community will suffer through the loss of many of its natural leaders, and through the chronic unemployment of a large fraction of its members. Unless invigorated from without, the whole community may throw up the sponge and wait apathetically for better things. In short, the problem of the Depressed Areas comes into existence: the problem, that is, of persistent unemployment concentrated in pockets up and down the country. It is a problem which cannot altogether be got rid of so long as areas specialise at all. But it might be made a less serious problem if the Government were to promote the diversification of industry in the Depressed Areas, and indeed in any area too narrowly

dependent upon a staple trade with an uncertain future. Some measures (e.g., the setting up of Trading Estates) have already been taken with just such an end in view. New industries have been attracted to the Depressed Areas by the offer of exemption from taxes, the grant of loans on favourable terms, the building of factories for rent at cheap rates, and so on. These new industries helped to reduce unemployment and put fresh life into the Areas; later, they will provide a second line of defence against renewed depression in the older industries.

The measures taken before the war were hardly drastic enough to work much improvement in the long-run prospects of the Depressed Areas. But they were a clear indication of the road which we are taking. Industry will not be left much longer to locate itself freely without regard to the over-riding interests of the community. We are moving rapidly towards government control of location. But if there is to be planned control there must also be that understanding of the forces localising each industry without which locational planning would be difficult and dangerous. We must understand the economic principles of industrial location and be able to apply them to particular industries and groups of industries.

## CHAPTER 6

## LARGE-SCALE PRODUCTION

SPECIALISATION, as Adam Smith pointed out, is limited by the size of the market. In a narrow market (e.g., in an isolated village), a worker must be able to turn his hand to many different jobs. The demand for his services in any one of these jobs will not be sufficient to keep him continuously occupied in it. In the same way, it will be impossible to use expensive machinery if it has to stand idle for long periods; or to introduce the elaborate organisation—in management, costing, research, marketing, and so on—which is indispensable in a modern business of any size. Production must be on a large scale before full use can be made either of machinery or of specialists and specialised departments. This is true of the establishment (e.g., a single factory), of the firm (which may own several factories), of the industry (which consists of a number of competing firms), and of industry as a whole. As each of these expands in size, new opportunities of specialisation, new types of machinery, and new kinds of organisation make it possible to achieve economies which were impossible so long as production was on a smaller scale.

The economies of large-scale production—called for short "economies of scale"—may be either "internal" or "external."

*Internal economies* are those which are open to a single factory or a single firm independently of the action of other firms. They

result from an increase in the output of the firm, and cannot be achieved unless output increases. They are not the result of inventions of any kind, but are due to the use of known methods of production which a small firm does not find worth while. They are also not to be confused with any bargaining or monopoly advantages which large firms may acquire. These advantages may swell the profits of large firms and incite small firms to grow bigger; but they do not (except indirectly) make large firms more *efficient* than small ones.

*External economies* are those which are shared in by a number of firms or industries when the scale of production in any industry or group of industries increases. They are not monopolised by a single firm when it grows in size, but are conferred on it when some *other* firm grows larger. Whenever an increase in the output of one firm has a favourable reaction on the efficiency of other firms, either in the same industry or in other industries, these firms enjoy the benefit of external economies.

## I—External Economies

The chief types of external economies have already been illustrated in our discussion of localisation.[1] There are:—

### 1. Economies of Concentration.

When a number of firms settle in a single neighbourhood they enable each other to derive mutual advantages through the training of skilled workmen, the provision of better transport facilities, the stimulation of improvements, and so on. Each employer has less difficulty in finding the kind of labour he wants, and can make use of special services—e.g., in merchanting his goods—which would not be available to scattered firms. In a country which has not yet been industrialised, economies of concentration will be of special importance. Each new firm, each new industry attracted to the country, will help to build up a machine-sense, and a common stock of knowledge and experience on which each succeeding manufacturer can draw. Every newcomer who settles alongside the pioneer firms will help the railways to reduce their transport charges and make it worth while to build better roads for all to use. All the social amenities characteristic of town life and beyond the means of a scattered population will begin to make their appearance.

### 2. Economies of Information.

In a large industry it becomes worth while to issue trade and technical publications to which everyone has access. Manufacturers are thus saved much independent research which, in a smaller industry, they might be forced to undertake

[1] See pages 57-8.

for themselves. It becomes possible also to set up research institutes—like, for example, the Department of Scientific and Industrial Research—which will carry out research work on behalf of individual firms and publish the results for any firm to use. Economies of information are clearly of much more importance when applied to industry in general than when confined to a single industry. Cheap newspapers and information bureaux, an efficient weather-forecasting service, adequate provision for research, and so on, are of enormous advantage to industry as a whole, but far too much for a single industry to provide unaided. Thus it very often falls to the government to bear the cost of information and research services, the benefits of which are shared throughout industry.

### 3. Economies of Disintegration.

The growth of an industry may foster the splitting off of some process which can be carried on more efficiently by a specialist firm or industry. On this further specialisation nothing need be added to what has been said above.[1] The economies which the specialist firms enjoy are internal economies and will be discussed presently.

What are external economies at one moment may be internal economies the next, if several firms unite with one another. To a village butcher, killing a few cattle a week, half of the animal is waste. But if the village grows into a thickly-populated urban district, with a large number of slaughterhouses, each of the local butchers will be able to dispose of this "waste" at a good price. Specialist firms will "make combs, buttons, knife and cane handles from the horns and hoofs; glue from parts of the bones and cartilage; pipe stems, chessmen, dice, artificial teeth, crochet needles or electrical bushings from the bones; pharmaceutical supplies from the glands, artists' brushes from the soft hair in the interior of the ears."[2] The butchers will benefit from external economies. But exactly the same economies might be enjoyed if the butchers were to amalgamate with the specialist firms to form a large meat-packing concern. The economies which, in the British livestock industry, are still largely external to the slaughterhouses could be made internal by the adoption of the methods of the Chicago packers.

External economies, therefore, are not different in kind from internal: what are external, and what are internal, economies depend simply upon what operations it is profitable to combine under a single management. In the steel industry, for example, external economies are likely to be of less importance than in the textile industries, because the processes of manufacture must be integrated in the interests of conservation of heat in steel-making, and separated in the interests of efficient management in the manufacture of textiles.

---

[1] See pages 57-8.

[2] S. H. Slichter: "Modern Economic Society," page 126.

Why the variety of processes or products of a single business should ever be limited at all—i.e., why the firm should not expand until it controls the industry, and why it should not then go on to absorb other industries as well—might seem, in view of the economy of large-scale production, to be a difficult problem. Why are all external economies not turned into internal economies through the amalgamation of all the firms in an industry or through the combination of all industry into one huge enterprise run, for example, by the State? These are problems to which we will return once the nature of internal economies has been made clear.

## II—Internal Economies

Internal economies can be grouped together under five headings: technical, managerial, commercial, financial, and risk-spreading.

### 1. Technical Economies.

Technical economies affect the size of the single establishment, rather than of the firm, which may own and operate several different establishments. An establishment may be anything between an ice-cream barrow and Ford's Rouge Plant at Detroit, covering over a thousand acres and employing 80,000 workers. In industries where technical economies of scale are great, the size of the typical establishment will tend to be correspondingly great; while in industries where size is of little technical advantage, and the methods of large plants can be easily duplicated on a smaller scale, the typical establishment will tend to remain small. In gold-mining, tin-plate manufacture, and the generation of hydro-electric power, the large plant uses methods of far greater efficiency than are open to smaller units; in those industries, therefore, large plants are the rule. In agriculture, on the other hand, it often makes little difference whether a farm is large or small. The dairy herd may be one of thirty or three hundred; the sheep flock, six hundred or fifty times as many. Apart, possibly, from the use of a tractor or a lorry on the largest farms, and the absence of milking machines from some of the smallest, farming methods will differ little from one farm to another.[1] If, therefore, we find many large dairy herds and sheep flocks, it will probably be because farmers are trying to take advantage of some economies of scale which are not purely technical (e.g., managerial economies).

Technical economies are of four kinds:—

**(a) Economies of Superior Technique.**—Many types of machinery cannot be reproduced at all on a smaller scale, so that a small plant

[1] A tractor becomes worth while on a "mixed" farm of about a hundred acres; milking machines are used when the dairy herd is as low as twenty. It is only in arable farming that farm machinery (e.g., the combine-harvester) gives the large unit a marked advantage over the small.

is driven either to instal machinery which it cannot keep in continuous use, or to put up with less efficient machinery of a different type. A large motor-works can use giant presses, which turn out all-steel tops, body-panels, and fenders in one piece; while in small factories a whole series of operations, more laborious and less mechanical, is necessary.[1] Only a large tin-plate works can make use of continuous strip rolling. Tabulating and calculating machines will be economical in a large office, but uneconomical in a small one. A rotary printing press, linotype machines, and so on, will be indispensable to a daily newspaper, but not to a small jobbing printer. Electrically-driven overhead cranes are in use in large engineering works and not in small. There are, in fact, few industries without some striking differences in technique between large and small establishments The small establishments generally make more use of small machines worked by hand-power; while the large establishments make more extensive use of large machines worked by steam or electricity. From a technical point of view, these modern sources of power are superior and this superiority is communicated, therefore, to large units of plant.

**(b) Economies of Increased Dimensions.**—Even when a machine can be duplicated on a smaller scale, there is often a purely mechanical advantage in using large machines. There is an economy, for example, when size diminishes losses by friction, or evaporation, or cooling. The carrying capacity of a ship increases in proportion to the cube of its dimensions; the resistance to its motion increases, roughly speaking, in proportion to the square of its dimensions. The power required to drive a given weight through the water is less, therefore, in a large than in a small ship. Again, if we double the dimensions of a water tank by building one twice as long, twice as broad, and twice as high, it will hold eight times as much water as before. But the walls of the tank will have increased only four times in area, and since the loss of heat by radiation depends upon the area of the walls, and not upon the volume of water in the tank, the rate at which the water will cool will be greatly reduced. A large boiler, therefore, will be more efficient than a small one. Similarly, a large furnace has great advantages over a small one. These advantages are of special importance in the chemical and metal industries where there is great scope for the conservation of heat through an increase in the scale of plant.

There is a further economy in the use of large machines. In spite of their greater output, they can generally be operated by a team no larger than is required for a much smaller machine of the same type. Whatever the size of a rolling mill, a steam shovel, a locomotive or a concentrator, the number of men required is much the same. "To turn a valve, to operate a switch, or to watch a gauge is no more

[1] That is, so long as there are no specialist firms from which the small factories can buy pressed steel parts.

arduous, though much more responsible, labour for a super-power station than for a tiny plant."[1]

Not only is the large machine cheaper to operate. It is also relatively cheaper to construct. An electric motor developing 20 h.p. does not require twice the material of a 10-h.p. motor, and takes little more labour to assemble. A double-decker 'bus is not twice as costly as a single-decker; the bodywork may be no stouter, and the engine, apart from a difference in horse-power, identical. In the building of a ship, a furnace, a factory, a telephone system, or a broadcasting station—to say nothing of the manufacture of pails and post-card albums—the size of the unit can generally be doubled without a doubling of labour and material costs.

It must not be supposed, however, that large mechanical units are *necessarily* more efficient than small ones; much depends upon what particular dimensions are increased, or what difficulties the increase in size is intended to overcome. The farmer who, impressed by the economies of increased dimensions, was silly enough to double the length, breadth, and depth of his plough, found that it needed not one but forty horses—with a large gang of men in attendance—to drag it through the earth.[2] When a ship is very big, the resistance to its motion is greatly increased by skin friction, and size becomes a handicap. The large unit, too, may forfeit its advantages if there is no labour capable of handling it properly, or no raw materials of the quality required. As an example of the labour problem, English maids are said to have difficulties with electric washing machines. As for the problem of suitable materials, Scottish blast furnaces have a very low capacity, chiefly because coke made from Scottish coal is too soft to stand a heavy burden of ore without becoming crushed and "choking" the blast.

(c) **Economies of Linked Processes.**—The operations of a single large machine are not likely to be the only ones undertaken in an establishment. There will be other operations, linked with those which the machine performs; and the more processes are linked together the larger, inevitably, will be the establishment. We must include amongst economies of scale, therefore, any economies resulting from this linking of processes.

These economies are not purely technical. The avoidance of risk, for example, is often an important motive to the linking of different stages of manufacture. A wireless factory may make its own screw nails to remove any danger of irregular or unpunctual supply. Motor manufacturers may build their own steel furnaces and casting shops because the quality of the materials which they buy is uncertain or difficult to test. Or it may happen that the firm has a reputation for a highly individual product, so that personal supervision at each stage of manufacture is essential.

[1] E. A. G. Robinson: "Structure of Competitive Industry," page 32.
[2] J. M. Clark: "Economics of Overhead Costs," page 116.

The linking of processes also leads to many technical economies. There is generally a saving in time and a saving in transport costs, since two departments of the same factory are closer together than two separate factories. It is mainly for this reason that the editing and the printing of newspapers are generally carried on in the same building, and that most factories have their own repair, testing, and box-making departments. In coal-mining, the washing and screening of the coal takes place at the pit-head, partly because, since coals of different grades are sold in different markets, a saving in transport costs is effected. In short, wherever a process would normally be localised in the vicinity of the raw materials, there is a strong tendency for it to be carried on within the establishment where the materials are prepared.

There is a saving, too, in fuel and power wherever the physical conjunction of two processes makes for the conservation of heat. The most striking example is that of steel-making. In an integrated plant, the pig-iron is taken to the steel furnaces in a molten condition, and the steel ingots, once they have solidified and have been re-heated in the soaking pits, can be rolled and re-rolled without being allowed to cool. The hot gases from the coke ovens and blast furnaces are used to heat the steel furnaces and soaking pits, and to generate the electricity which works the rolls in the rolling mills and the blowers in the blast furnaces. Thus the consumption of coal per ton of finished steel can be reduced from the 1913 average of over three tons to less than half that amount in a modern plant like Lysaght's works at Scunthorpe.

A further saving arises when waste materials become available in quantities sufficient to justify their manufacture into by-products. There is here a double incentive to the building of large establishments, for not only must a new "salvaging" process be added, but the parent process must be enlarged until waste materials are available in bulk. The classical example of the large-scale manufacture of by-products, and the savings which it makes possible, is the meat-packing industry, in which the value of by-products may come to as much as 40 per cent. of the value of the meat. Many examples are provided also by the largest establishment in the world—Ford's Rouge Plant at Detroit. For instance, the blast furnace slag, which in a small plant would probably be valueless, is sold as road-making material, or granulated and used in making Portland cement. In the stack of the cement plant, special filtering equipment collects about 40 tons of fine particles a day, and these, when re-processed, yield 252 barrels of finished cement. Waste paper, lunch wrapping, and similar materials are salvaged and made into millions of feet of cardboard in the Rouge paper mill. Valuable chemicals and charcoal briquets are made from the hundreds of tons of chips and shavings reclaimed from lumber mill operations. Thousands of gallons of oil are recovered daily by running steel turnings through centrifugal wringers, cleaning the oil (which adheres to the turnings) of grit and foreign matter, and returning it to the

factory to be used again. Few of these salvaging operations would be worth while in a smaller plant—one, for example, with no paper mill and no steel turnings.

Once several processes are linked within the same plant, it becomes necessary to *balance* them against one another. If the plant is to operate with maximum efficiency, the scale on which each process is conducted must not only exhaust the economies possible at that stage, but must be adapted to the demands of the succeeding stage. One machine, when working to capacity, may be able to handle 100 units of output; the capacity of a second machine may be 250 units, of a third 300, and of a fourth 2,000. If each of these machines is to be kept in gear with the others, it will be necessary to use more than one model of each machine, and to produce more than 2,000 units of output. The ideal solution would be to scale up output to 6,000 units (i.e., the L.C.M. of the capacity outputs of each machine) and use 60 machines of the first type, 24 of the second, 20 of the third, and 3 of the fourth. In the real world, however, capacity outputs do not fit together so neatly as in this example, and the L.C.M. would often be impossibly high. Some possible economies, therefore, have to be sacrificed by overdriving one machine and running another below capacity. On the other hand, the risk of disorganisation through breakdowns—especially at bottlenecks like our fourth machine—will tend to push normal output beyond the L.C.M., since the operation of a second unit will act as a safeguard against complete stoppage all long the line. Similarly, if routine repairs (e.g., the re-lining of blast furnaces) can be carried out only when a large unit of plant stands idle, several units of the same type will usually be operated together.

**(d) Economies of Increased Specialisation.**—A further advantage of large plants is the scope which they offer for increased specialisation. Division of labour is carried further than in small plants, since workers can be kept fully occupied at a comparatively narrow range of tasks. In a small garage, a man may be mechanic, taxi-driver, and petrol attendant. In a small office—e.g., in accountancy, banking, or law—a clerk has generally more varied duties than in a large one. But in large and small factories most jobs are much alike. Most of the benefits of division of labour are enjoyed even in comparatively small plants, and in garages, offices, etc., the balance of advantage is not always on the side of the large specialised unit.[1]

Technical factors cannot limit the growth of a firm; they create economies of large-scale production, never diseconomies. At a certain scale of output, no improvements in technique will result from further growth. But neither will there be any deterioration. The optimum, or most efficient, technical unit can be duplicated and multiplied without loss of efficiency. If, therefore, there are other

[1] This conclusion is not intended to apply to management and supervision, which are discussed in the next section.

savings to be made—in management, in finance, in marketing, and so on—there will be every reason why the establishment—and, still more, the *firm*—should gain in efficiency by growing beyond the technical optimum, and no reason why technique should set limits to expansion. We must look in other directions for obstacles to growth.

### 2. Managerial Economies.

Managerial economies may be effected either by increasing the size of an establishment or by grouping a number of establishments under one management. Both methods of expansion create opportunities for specialisation on the part of the managing staff. This specialisation is twofold:—

**(a) Delegation of Detail.**—In a small business, the owner spends much of his time going to and fro in his shirt sleeves, superintending and assisting with the work of his factory. He has probably a thorough grasp of every side of the business and a strong hold on the loyalty of his workers. He may enjoy being worker, foreman, and manager combined, and have no wish to change. But if he is an able organiser, he will produce at lower cost—and incidentally make a larger profit—by expanding his business and delegating routine and details to subordinates. This will leave him free to concentrate on the more difficult work of organisation. He will have time to make a full study of special problems, keep abreast of changes in technique, make influential "contacts," and form a plan of campaign on the basis of adequate information. But all this will only be possible in a large business, where his time is not taken up with what a lower-paid worker could do equally well, and where his powers of foresight, initiative, and judgment are spread over the maximum volume of output to which they will stretch. Mr. Ford and Lord Nuffield would be wasted on a poultry-farm or the local cinema; they might even fail disastrously. But they are so far from being wasted on the Ford and Morris Motor Companies, where their organising ability is given full scope, that even the payment of a million-dollar salary might still leave these companies the gainers.

**(b) Functional Specialisation.**—Division of labour can also take place, so to speak, horizontally. The work of management is split up through the creation of separate departments, each of which is put in charge of an expert administrator. One specialist becomes responsible for the organisation of production, one for sales, one for transport, still another for the maintenance of machinery and buildings. There will be other departments engaged in costing and estimating, research, the hiring of workers, the purchase of materials, the preparation of new designs, the keeping of accounts, the arranging of finance, and so on. Each large department may be further sub-divided and its duties distributed between a large number of specialists. The sales department, for instance, has not only to make and enter up sales satisfactorily; it has also to keep customers

in good humour, find out where sales resistance is weak, guess what people will want to buy, and persuade them to buy it. In a small business, these are minor duties which the owner-manager must take in his stride; the average farmer, for instance, does not lose much sleep over them. But in a large modern business nothing short of the appointment of a public relations officer, and the setting up of market research and advertising departments, will get them done.

The same division of functions takes place on the board of directors. The job of the directors is to guide the general policy of the business, much as Cabinet Ministers are supposed to govern the policy of government departments, without undertaking the detailed work of administration. One director may make himself responsible for finance, another for marketing, and a third for engineering problems. Thus the board of directors may be turned into an all-star team, in which are knit together widely different talents and experience. Each of the directors will have his blind spot; and, if he were in complete control, might soon turn upside down the organisation of departments unfamiliar to him. But in a large business, with any luck, he can be suitably chastened and chained. There will be enough to keep him busy in Department A, for which he has a flair, and no need to let him loose on Department B, in which he is a comparative duffer.

The appointment of specialists, either as directors or as managers, and the setting up of specialised departments, is economical only in a large business. A research chemist, an accountant, or a works manager may draw a salary of £1,000 in a large concern and be cheap at the price; but he is much too expensive for a small one. A large business can afford a separate staffing department which will test applicants for vacancies and assign them to posts for which they are suited; but in smaller businesses, where vacancies are few, recruitment is haphazard and costly, much time being lost in the training of misfits. Similarly with research: a small firm cannot possibly spend large sums on its research laboratories, but a large one may spend several million pounds a year. The American Telephone and Telegraph Company "has several thousand people at its headquarters and laboratories who do nothing but make new inventions and develop new methods of operation."[1] But as this company has been known to spend over £100,000,000 in one year on extensions and improvements alone, its research budget is a mere trifle in comparison with the economies to which the research may lead. General Motors Corporation has a testing ground with roads of every description—mud, dirt, sand, gravel, and concrete. There are hills with gradients up to 1 in 4, a four-mile speed track, and a complete service station. Yet the cost of operating the testing ground is no more than a few cents per car sold by the company.[2] The expenditure on research by a large British company like

[1] S. H. Slichter, *op. cit.*, page 13.

[2] S. H. Slichter, *op. cit.*, page 129.

Imperial Chemical Industries is just as enormous in comparison with what smaller companies can afford and just as insignificant a fraction of the company's total budget. The full benefits of a research, or any similar, department can be enjoyed only by very large enterprises.

Economies in management, then, like technical economies, can be realised only when production is on a large scale. But they cease to be realised once firms grow beyond a certain size. Instead, they give way to managerial *dis*economies. The larger the firm, the more complicated become the problems of management. Each department has to be kept in step with all the others; its programmes must be co-ordinated, its needs supplied, its shortcomings detected, its grievances dealt with. A steady flow of materials must flow through the plant so that work is not held up for lack of supplies. The organisation of production must constantly be altered to deal with rush orders or loss of markets, breakdowns, or failure to obtain delivery of raw materials. The strain of planning and co-ordinating is enormous even in a comparatively small business, where decisions can be taken at top speed and reversed almost as fast. In a big, departmentalised concern the breaking-point is often reached. The man on the spot will shirk the responsibility of a quick decision, and refer the matter to his superior or to some other department. The superior (or the department) consulted will refer the matter back still further, or take refuge in routine and red tape. And quite naturally, too. If subordinates are picked to attend to details while their superiors use their brains from afar, a hitch from time to time is inevitable. If one department is given specialised duties, an occasional lack of liaison with other departments, or an excess of zeal not to usurp their duties is only to be expected. The drawing office may go without pencils until the storekeeper supplies them; the sales department may run out-of-date vans because the managing director has never seen them; the market research department may be always at loggerheads with the designers. With responsibility diffused and uncertain, the manœuvres for authority between rival officers and departments may go far to offset the gains of specialised management.

The same loss of flexibility, and the same danger of personal rivalries, may be felt in the framing of high policy by the board of directors. An all-star team made up of captains of industry may not combine too well. There is the difficulty of getting agreemen between strong personalities. There is the temptation to caution because so much hangs on a decision, and because to change or reverse it later may be next to impossible. There is the loss of drive because each of the directors is a hired official, not an owner-manager, playing for his own hand and free to pick his own lieutenants. The larger the business, the more unwieldy and bureaucratic it becomes. The manager of a small firm is in touch with every side of the business; he can consult his subordinates and make up his mind quickly. But in a large firm there may be endless delay, during which files and

memoranda accumulate, before the board can take action. High specialisation, in management as in everything else, means a looseness of contact between the specialists—between the chairman of the company and the office boy, the engineer and the economic adviser, the advertising and the research departments. These loose contacts breed misunderstandings, oversights, and delays. Blunders are easily made; and the cost of these blunders is greater, the greater the size of the firm. If the firm, in growing, pushes specialisation too far, these blunders will ultimately be so damaging to efficiency that they will more than offset the advantages of further growth.

At what stage in a firm's growth this will happen varies from industry to industry and from firm to firm. Managerial diseconomies set elastic limits only. Why this should be so, why the small firm should survive more readily in some industries than in others, we shall see once the other economies and diseconomies of scale have been discussed.

### 3. Commercial Economies.

Commercial economics in the purchase of materials and sale of goods may be just as important as economies in manufacture. In the eastern counties of England, for example, a recent inquiry showed that smallholders were paying 10 per cent. more for their feeding-stuffs and 20 per cent. more for their fertilisers than "large-scale" (300 to 500 acres) farmers. Smallholders were also in receipt of lower prices for their grain. Had they bought and sold on the same terms as the large farmers their income would have been greater by £15 a year—by no means an insignificant fraction of the family income.[1]

In other industries the possibility of making substantial economies in buying and selling is equally great. Raw materials may come to 50 per cent. or more of the cost of an article. Nearly 40 per cent. of the cost of a newspaper, 60 per cent. of the cost of bread, and 75 per cent. of the cost of cotton yarn goes on raw materials. Selling costs may reach even higher proportions. The cost of labour and materials in patent medicines is generally negligible in comparison with the price. Half of the price of milk goes to meet the costs of the distributor. In furniture, jewelry, and luxury goods of all kinds, a distributor's margin of 50 per cent. and upwards is common. Clearly, a slight saving in the cost of distribution may be more valuable than a substantial cut in manufacturing costs. What we have to decide, however, is not whether there is room for economy, but whether there are more opportunities to make such economies in a large than in a small business.

First of all, the large buyer is able to obtain preferential treatment. He is quoted the lowest prices and offered discounts and rebates.

[1] A. W. M. Kitchin: "Smallholdings and the Agricultural Structure," *Economic Journal*, 1934, page 668. In 1932 the family income of the 200 smallholdings analysed was £72 (excluding wheat deficiency payments from the government).

He can obtain freight concessions from railway and road transport companies, liberal credit from the banks, and prompt delivery, careful attention, and special facilities from manufacturers. Even when others go short of materials, he is still able to obtain them. This preferential treatment is often the fruit of bargaining strength. Prices must be "cut to the bone" because the large buyer might enter into competition with the firms from which he buys. Thus Ford is able to buy steel sheets 5 per cent. or more below the market price, partly because he operates steel furnaces as a yardstick with which to measure costs.[1] In the same way, firms which might run their own fleet of motor vans and lorries are in a strong position to demand freight concessions. Or if the market is narrow, and the large buyer has few important competitors, he will be able to bring pressure on the firms which supply him by threatening to withdraw his custom and influence.

The bargaining advantages of large businesses make them profitable. But they do not make them *efficient*. A big firm may put money in its pocket by using its bargaining strength to beat down the price of materials. But it does so only at the expense of some other firm. The mere transfer of profit from one firm to another works no more improvement from the social point of view than the operations of gamblers or pickpockets. It is not a true economy at all. Similarly a big firm may find itself in a position of monopoly and mark up the price of its products, or sell an inferior quality at the old price. Here, too, there is no economy but only the exploitation of a monopoly advantage. If, therefore, we ask, not whether large firms pay but whether they are efficient—and that is the only question under discussion in this chapter—we must leave the fruits of bargaining strength out of account. Preferential treatment to large-scale buyers is a sign of economy in marketing only when it is offered by the seller in consequence of a reduction in his costs.

Such a reduction will normally take place when a large order enables plant to be run to capacity and without constant adjustment and frequent spoilage. A truckload will not cost twice as much to transport as two half-filled trucks. Steel sections rolled to one specification will cost less than an assortment of sections of the same weight, since the work will not be held up by the need to change the rolls. A large order which gives a factory steady work for six months is obviously more satisfactory than a series of small orders which provide work intermittently. In these and similar cases, better terms can be quoted for large orders than for small, and the firms which are big enough to place large orders make corresponding savings.

A reduction in costs also takes place because large orders enable the sales staff to be used to capacity. If customers double their orders it will not be necessary to double the sales staff. A packet

[1] And also as a safeguard against failure of supplies (e.g., because of a strike).

of twenty cigarettes does not take twice as long to sell as a packet of ten. A dairyman takes very little longer if he has to deliver eggs and morning rolls with the milk. A traveller for one variety of tinned foods can book an order for fifty-seven with little extra trouble. There is more here than a saving in time. A manufacturer who has built up a reputation, or a distributor who has a good clientèle, for the sale of one product, can sell others on the footing of his reputation or connection. If motorists buy your tyres they may buy your golf-balls, too. Every commodity which a firm sells is an advertisement for every other which it is trying to market. Manufacturers, therefore, may be induced to expand their business by taking up new lines, and so using their sales organisation and their sales advantages to the full. At the same time, they will have a further motive for offering better terms to firms placing large orders, and these firms, in their turn, will be tempted to expand.

The large firm, then, enjoys preferential treatment, partly because of the economies of large orders. It also enjoys other advantages in marketing. It can employ expert buyers, skilled in the selection and blending of materials. It can submit its materials to strict tests, grading them so that the final product is of uniform quality, and reducing the risk of waste from faulty materials. It can generally bide its time in buying and selling, and refrain from making purchases or sales when the market is unfavourable.[1] Its customers will have the advantage of a better service of repairs and replacements than those who have bought from smaller firms. In making their purchases, they will have a wider range of choice, out of a greater variety of stocks, and will frequently be able to obtain immediate delivery from stock of goods which a smaller firm would require time to manufacture.

### 4. Financial Economies.

The large firm has also many financial advantages. It has a wider reputation and more influence amongst those who have money to lend or invest. It can borrow from bankers upon better security and raise capital more readily through the issue of shares and debentures than a small firm. There is a wide and regular market for these shares, so that shareholders can realise their capital without any of the trouble to which they would be put in a small private concern. Thus the cost of obtaining credit or of raising fresh capital is lower for a large than for a small firm.

[1] On the other hand, it must often make contracts for large quantities so as to allow of steady production. These large orders, placed openly, may force up the price against the company—all the more if the number of producers who can guarantee delivery is no more than two or three. A small competitor is in a much better position to "shop around" the market, obtaining secret concessions from the published price and buying odd lots when the opportunity offers. For an example from the glucose trade, see A. S. Dewing: "Financing of Public Corporations," page 760.

This difference in the cost of borrowing might seem to be proof of a bargaining advantage rather than of true economy. For a large business to be granted easy credit because it has several bankers on its board of directors is graft and not economy. For a shareholder to invest in a company whose shares are quoted in the financial pages of the newspapers rather than in one of which he has never heard, does not mean that the first company will make better use of his money than the second. We do not look on a company which hires its labour very cheaply as an efficient company (although it may be); why then should a company hiring its capital cheaply be given that title?

All this is true. But it is incontestable also that the lending of money in large quantities, like the bulk purchase of materials, is less costly than the lending of money in small quantities. A system of large-scale production allows capital to be raised with more convenience to the lender than a system where production is in the hands of a large number of small firms. This convenience to the lender—for instance, the greater ease with which he can recover his capital—is an economy of scale exactly like the convenience of supplying a large order rather than a series of small ones. The convenience of large-scale borrowing will be discussed in more detail in the next chapter.

### 5. Risk-bearing Economies.

Large firms are often less exposed to risk than small ones. First of all, they are able to eliminate risks by grouping them. We can often predict what will happen *on the average* with fair certainty when the *individual* items defy prediction altogether. Thus we can say with some confidence that 6,000 people will commit suicide in Britain next year, but we cannot possibly say who they will be. Similarly with births and deaths, motor accidents, burglaries, fires, and all the contingencies against which we can insure. The larger the number of instances, the less the risk of errors in judgment. In a large business, therefore, where the same operation is more often repeated than in a small one, uncertainty can be reduced. "To the owner of a cow the loss at calving time is uncertain, while to the owner of a great herd this loss appears as a regular percentage that can be computed and allowed for. . . . Again, in a small refinery the possibility of over-doing a batch of oil or sugar may be a source of serious uncertainty, while in a large refinery the law of the average prevails."[1]

The grouping of risks is of particular importance in finance. A large bank or insurance company, for example, tends to be more stable than a small one because it is less likely to be overwhelmed by a purely local catastrophe. The collective investments of a large bank are far more secure than the investments of each of the branches.

[1] E. A. Ross: "Uncertainty as a Factor in Production," quoted in C. O. Hardy: "Risk and Risk-bearing," page 19.

Similarly the insurance benefits payable by a large company vary less from year to year and can be more accurately foreseen than the benefits payable by a small company. Through their greater stability, large financial enterprises are able to command the confidence both of their customers and of their investors. Thus they are in a position to quote attractive terms for their services and to raise their capital comparatively cheaply.

When similar risks are grouped, the law of averages applies and uncertainty is reduced. When dissimilar risks are grouped, the advantage is less definite, but far from negligible. The spreading of risks is, indeed, one of the main preoccupations of modern industry. Firms may seek to spread their risks by diversifying their output, or markets, or sources of supply, or processes of manufacture. Such diversification makes the firm less vulnerable to sudden changes and allows it to remain a going concern where smaller, less diversified, concerns would be forced to give up business.

**(a) Diversification of Output.**—Where a firm is likely to be injured by a decline in the demand for its product, it will be likely to look round for other products which can be easily manufactured simultaneously with the first. In this way, it may be able to level out a slump in one product against a boom in another. Or if something happens to interrupt the manufacture of one, it will still be possible to carry on with the manufacture of the others. One of the main reasons why mixed farming has such a hold on British agriculture, for example, is that it is unusual for all branches of farming to be equally depressed, so that the farmer who is selling several different products can make ends meet by setting the profits from one against the losses on the others. If the depression lasts a long time—as agricultural depressions do—he can cut down the production of his biggest losers and switch to the more profitable branches of farming (e.g., from stock-raising to dairy-farming, or from wheat-growing to barley and oats).[1] Similarly one may find a fish-and-chip shop selling ice-cream, or a refrigerating plant dovetailed with a skating rink. One firm will manufacture mustard and barley water; another, electric equipment and wireless sets. A motor-car manufacturer will be careful to offer several different models to improve his chances of catching the public fancy for one of them. The more fickle the demand, either from one season to another, or from year to year, the stronger will be the tendency to spread risks and steady production by diversifying output.

**(b) Diversification of Markets.**—Even when only one product is manufactured, increased sales, if there are to be a larger number of markets than before, may reduce the danger of fluctuations in

[1] There are also many technical advantages (e.g., the elimination of bare fallows by an alternation of fouling and cleaning crops, and the more even distribution of labour requirements throughout the year).

demand. The sale of electricity for all purposes is less variable than the sale to industrial consumers only; the total sale to domestic consumers is more stable than the sale for heating, lighting, or cooking only. It pays distributors, therefore, to develop the domestic market; or, if they already supply electricity for lighting, to encourage consumers to use electric radiators, refrigerators, and cookers—to say nothing of radios, razors, and hair-wavers, all of which are, or should be, used at different times of day, and sum up to give a comparatively stable demand. In the same way, a railway which is largely dependent upon mineral traffic may find it worth while to develop suburban passenger transport. A firm which sells a new product in a local market may seek to make the demand broader and steadier by advertising it on a national scale.

(c) **Diversification of Sources of Supply, and (d) of Processes of manufacture.**—Here the same principle applies. Large firms may be better able to maintain output when some particular source of raw materials is cut off (e.g., by a strike), or when some particular process of manufacture becomes uneconomical or impossible. If, for instance, they use steam power as well as electricity bought from a central distributing agency (the "grid"), a failure in the supply of electricity, or a rise in the price of coal, will not hold up production altogether. A large 'bus company, owning oil-driven and petrol-driven 'buses and trolley-'buses worked by electricity, will be less vulnerable than a smaller company owning only one kind of 'bus. Similarly, a large business is likely to draw its supplies of raw materials from a wider area than a small one, and will be less exposed, therefore, to the danger of having its supplies cut off altogether.

It may seem strange to think of the spreading of risk as an economy; but if we mean by economy, "making the best use of resources, including human judgment," the term is quite properly applied to the spreading of risks that are known to be great. A firm which goes on producing regardless of risks may turn out goods at very low cost for a time, but if it is left high and dry by a change which might have been foreseen, and which other firms foresaw, its costs in the long run will work out above, and not below, theirs. For it will have sunk capital in an organisation which is put out of action by a change which might reasonably have been predicted. It will have led resources, to put it vulgarly, up the garden path, and committed them to uses in which they are of little further value. It would have been more *economical* to spread risks at the start by sacrificing some of the advantages of specialisation for the safeguards of diversification.[1]

[1] This line of argument must not, of course, be pushed too far. There are limits to the advantages of spreading risks. What these are depends upon what view it is "reasonable" to take of the future, and how strongly one may "reasonably" hold that view.

Risk-spreading economies only make for large-scale production if they do not greatly complicate the business of management. But this complication is hard to avoid when products, processes, and markets are diversified.[1] Risk-spreading and managerial economies, therefore, very often pull against one another. Sometimes the net result is in favour of large firms; when, for example, two products with different seasonal peaks can be dovetailed under one management. At other times the spreading of risks can be accomplished only in comparatively small firms. A mixed farm, for instance, requires far more supervision than a wheat farm in the prairies, and is much smaller, in spite of the larger number of things which it produces.[2]

Risk-spreading and technical economies are also in conflict with one another. A firm cannot obtain the full benefit of the technical advantages of large-scale production unless it sinks a great deal of capital in machinery and plant. If the plant can be run to capacity, so that it produces steadily on the scale for which it was designed, technical economies can be exploited up to the hilt. But if there is a risk of discontinuous operation, small firms using less expensive machinery, or less elaborately organised, may come into their own. In the steel industry, for example, it is very often the relatively small plants which survive an industrial depression most successfully. They are greatly inferior, from the technical point of view, to large integrated plants designed to handle enormous outputs. But they have the advantage of lower standing charges on their machinery and plant, and have less difficulty, therefore, in making enough profit on a low output to cover these charges. They can also make savings more readily on operating costs if output has to be reduced below capacity. In a large integrated plant, there is a nice balance between the blast furnaces and coking plant, the steel furnaces and the rolling mill. But this balance is completely upset if the plant has to be run below the level of output for which it was designed. The heat supplied by the coking plant, or consumed in the soaking pit, cannot be halved to match a halving in the output of steel. If blast furnace gas is used to generate electricity, it will be impossible to damp down the furnaces and keep the rolling-mill in operation. The team of men who work the plant at full capacity will be almost as large as the team required for an output 20 per cent. lower. In a small firm operating costs are more flexible. A firm owning no blast furnaces, for example, will find it easier to substitute steel scrap for pig-iron in its furnaces whenever scrap is comparatively cheap. It will also be more adaptable

[1] "The more restricted the variety of the processes undertaken by a single business, . . . the simpler is the task of direction. The more simple the work of direction, the larger is the volume of output which can be efficiently controlled by a single mind. And the larger the volume of output, the greater is the scope for the major economies of modern technique." F. Lavington: "Technical Influences on Vertical Integration," *Economica*, 1927, pages 27-8.

[2] Smaller, judged by the total value of what it produces, and not simply in point of area.

to quick *changes* in output. The costs of damping down and relighting furnaces, or of stopping and starting again in any other part of the works, are high in a large plant, where regular operation at a constant level of output is indispensable to efficiency. In a small steelworks changes in output are less disorganising and can be made at less cost. Thus the risk of change, like the risk of operating below capacity, tends to halt the growth of the firm below the technical optimum.

The spreading of risks through an increase in sales will normally also reduce marketing costs. This double economy in marketing and risk-spreading is always a temptation to firms to put new products on the market, especially when they have their own sales organisation. A familiar example is the farmer who takes to feeding chickens as a side-line and sells the eggs from his milk-cart. But risk-spreading, as we have seen, does not always make for an increase in output and sales. In mixed farming and steel-making, for example, risks are on the side of the small firm. Some marketing economies may have to be sacrificed, therefore, to permit of the spreading of risks by the small farm and steelworks. The large wheat farmer who changes over to mixed farming will find that his marketing costs have risen.

A full account of how industrial organisation can be adapted so as to make the fullest possible use of economies of scale is beyond the scope of this book. For the present, it is enough to grasp the aim of a firm in trying to blend, or reconcile, these economies with one another. If firms habitually produced at the lowest possible cost, they would expand until they reached some optimum size at which there was an exact balance between the possible economies to be reaped by further expansion and the diseconomies associated with expansion. Firms producing on a scale smaller than this (i.e., of less than optimum size) would be incurring higher costs, e.g., by sacrificing the use of expensive machinery without any compensating advantage. Firms of more than optimum size would also be incurring higher costs, e.g., because the organisation was cumbersome and beyond the powers of its directors to manage properly.

From the social point of view, it would be a great gain if firms automatically grew to their optimum size, and produced on the scale at which costs were lowest.[1] Unfortunately, this does not happen. Most firms neither do, nor try to, reach their optimum size. Their aim is to make profits, not to produce with the greatest possible efficiency. It is not cost, but the excess of price over cost, that controls the size to which firms grow. They may stop short of the optimum rather than force a larger output on the market at a lower price; or push beyond the optimum for the sake of competitive strength.

We can now bring together our results. A wide variety of economies favour production on a large scale, but there are a number of obstacles to growth. Firms can have access to economies of scale only if

[1] This statement is suitably qualified later. See page 136.

they are of a certain minimum size. In growing towards this size, however, firms meet with increasing difficulty in maintaining efficient management and in adapting themselves to the uncertainties which they have to face. If their sole aim were to produce at minimum cost, they would tend to reach an optimum size at which unexhausted economies of scale were just offset by potential diseconomies. In fact, however, their growth depends upon two additional factors (not yet analysed)—market resistances, and the desire for competitive strength.

## CHAPTER 7

## SMALL-SCALE PRODUCTION

IN the last chapter our attention was concentrated on the advantages of large-scale production, while the forces making for small-scale production were touched on very briefly. In this chapter we shall look at the other side of the picture and try to discover why small-scale production still holds its own. The general line of reasoning which we shall follow is that small firms will find it easiest to survive in industries where the optimum size of firm (as modified by market resistances and the desire for competitive strength) is small. In other industries where economies of scale—and especially technical economies—are great, small firms stand little chance. Small steel firms, for example, are wiped out while small shops flourish. Unless there are obstacles to growth which come into play early, and are not broken down by the overwhelming technical superiority of large-scale units, small firms cannot readily survive. These obstacles to growth can be grouped under the same headings as economies of scale, except that, as we have seen, there are no technical obstacles.[1] Obstacles associated with risk have already been discussed.[2]

### 1. Managerial Obstacles.

First of all, the range and complexity of the problems of management are greater in large than in small firms. In any industry, therefore, where there is need of constant supervision and rapid decision—where each firm must use a great deal of management, or "decision-taking," per unit of product—small firms will predominate. Their advantages lie in the absence of divided responsibility, in the attention which they can give to detail, and in their ability to cater exactly for the wants of their customers. In retail shopkeeping, in farming, and in trades where the influence of fashion is great, or a very high quality of product is demanded, the small firm has no difficulty in holding its ground. But in trades that can be reduced largely to routine—the railways, the post office,

[1] See above, pages 68-9.

[2] See above, page 78.

etc.—or in which the range of problems is narrow—in cotton-spinning as compared with cotton-weaving,[1] or in the worsted as compared with the woollen trades[2]—the advantage lies with the larger firm.

Given the range of problems in an industry, the large firm may be assisted by the development of new forms of organisation and new methods of business management. The rise of cost accounting, for example, makes it easier to cope with the problem of tracing waste, and allows comparisons to be drawn between the efficiency of different departments and factories. The joint-stock system provides the large firm with facilities for raising enormous amounts of capital, and smooths the path of able men as they try to expand their business to the limits of their organising ability. The spread of education and the growth of business morality makes it easier to delegate tasks to subordinates. Thus in industries where advantage can be taken of cost accounting, scientific management, up-to-date office equipment, and so on, even complex problems may be handled satisfactorily. But in other industries the complexities will only be made more complex. Small firms which avoid efficiency experts have often good cause to congratulate themselves. And more cause in some industries than in others. Modern business methods have done little to increase the size of farm which one man can manage; but they have greatly increased the size of wireless factory.

The managerial obstacle to growth is, in the last resort, a personal one. There is a limit to the range and complexity of the problems which a business man, however able and however ably served, can tackle satisfactorily. The burden of detail becomes enormous, particularly in a business made up of scattered plants faced with their own special problems. If there is little devolution of responsibility to departmental and branch managers, they are deprived of initiative and degenerate into mere cogs in the machine. If, on the other hand, responsibility is delegated, subordinates of outstanding ability must be appointed and must be handsomely paid. These subordinates may, and frequently do, launch out on their own in competition with their former employers. The zeal and energy which they then show for their own interests may be out of all proportion to their previous efforts on behalf of their employers, for as salaried officers they have much less at stake, and are much less inclined to make sacrifices. Even when loyal and responsible subordinates can be hired fairly cheaply, a business man is still faced with the difficulty of exercising effective control over them. But he is unlikely to be able to manage more than four or five departmental heads without causing endless delays or

[1] Cotton yarns are spun by standard methods to standard specifications in weaving, on the other hand, the mills generally produce a fairly wide range of product, differing in quality and pattern.

[2] The woollen trades call for much more skill in blending from a great variety of raw materials.

sacrificing his authority and becoming no more than a rubber stamp.[1] If a way out is sought through the substitution of a board of directors for a single "boss," the range of problems with which the board can deal may be increased, but again there will be an inevitable limit.

The more plentiful is organising ability, the weaker is the managerial obstacle. That outstanding organising ability is not plentiful, and that the managerial obstacle is strong, is suggested by the fact that a very large proportion of our giant businesses have been built up in their own lifetime by a few exceptional men. We can point to Ford and Carnegie, Austin and Leverhulme as the creators of the businesses with which their names are associated, and find an explanation of the size of these businesses in the organising capacity of their founders.[2] Napoleon's dictum, "How rare *men* are!" applies as much to industrial as to military organisation, and points to one further explanation of the survival of small firms.

But even if we were all first-rate organisers, the small business would not disappear. For the managerial obstacle depends also on our inclinations. From a mixture of motives—from a love of independence and uncertainty, from pride or ambition, or the urge to create[3]—men may prefer to run a small business of their own rather than act as subordinates at a higher rate of pay. Their workers, too, may prefer to remain with an employer who knows their habits and history rather than move to a larger business where there can be no personal contact between employers and employed. Given such a frame of mind, the survival of small firms may argue not so much a deficiency as a superabundance of talent. The development in *all* business men of greater organising capacity need not make for the extinction of the small firm. What *will* strengthen the big firm against the small will be the development of a *cleavage* in ability and temperament between the "captains of industry" and the "smaller fry." Where the "captains" are extremely able and are completely wrapped up in the task of carving out an empire for themselves in the world of business, it will be hard to stop them from expanding. The smaller firms, managed by less competent men, who seek a livelihood rather than power, will be tempted to throw in their hand and sell out or knuckle under to their more aggressive competitors. If there is a gulf set between the really able and the second-rate, and if the second-rate are disinclined to hold on to their independence while the really able are determined to expand, then small firms will rapidly disappear.

---

[1] Compare P. S. Florence: "Logic of Industrial Organisation," pages 121-2.

[2] Since it is easier to find men to manage large businesses *once they have been built up*, a steady trend towards bigger businesses is no proof that organising ability is becoming more abundant.

[3] One motive of some importance is the desire of the owner to leave to his son an occupation as head of his business and not just the capital which could be realised by selling out to a combine.

But in the real world there is no such gulf. The "captains" die off, or become "money-logged,"[1] and their businesses come into the hands of less able men. The "captains" of the second generation have to fight their way up from small beginnings, in charge of firms that are steadily expanding. We do not find vigour of management only in large businesses and incompetence or half-heartedness only in the small ones. The firms of each industry, to use Marshall's famous analogy,[2] are like the trees of the forest, some growing and becoming stronger every year, some losing vitality and giving place to others. There is a process of growth and decline in which firms of many different sizes compete for survival.

From the point of view of management, therefore, small firms may survive, for one of three reasons. They may be large firms in the making; they may be capably managed by men who are indifferent to the possibilities of growth; or they may be managed by men who lack the judgment, experience, and organising ability necessary for growth. Similarly, large businesses may be dwindling towards the level of their management; or they may be growing because of the enterprise of the management; or they may be changing little in size and be managed with more than average competence.

## 2. Market Obstacles.

Where the market is too narrow to permit of large-scale manufacture, the firm necessarily remains small. A firm might reduce its costs by doubling its output, but if the firm could not dispose of the extra output, there would be no sense in producing it. It is only when the market is big enough that no economies of scale need to be sacrificed and firms can expand freely towards the optimum.

The market may be limited in various ways.

(a) **Geographical.**—First, it may be limited by distance, as, for example, the market of the local baker is limited. Rather than incur high transport costs (including waste in transit) by supplying distant customers, firms may prefer to produce on a small scale for local needs only. Whenever firms are distributing their products over a thinly-populated district, expansion will involve heavy outlays on marketing to set against any savings in the cost of manufacture, and the firms will remain small. This is true also of sources of supply. If the raw materials of an industry are widely scattered, and are expensive to assemble at a central site, production will tend to be carried on in scattered factories drawing on local supplies.

[1] Compare, for example, Northcliffe: "The danger in businesses like ours is that they get settled down. People running them get comfortably off, 'money-logged,' and have no incentive to new enterprises. They forget that if you do not go forward you must go back." Tom Clarke: "My Northcliffe Diary," pages 259-60.

[2] Marshall: "Principles of Economics," page 315.

The milling of timber, grain, and minerals illustrates this tendency, while the furniture trade and the light castings industry illustrate the first.

The obstacle of high transport costs is all the greater when markets and sources of supply are not only diffused but also overlap. If production is carried on by small firms up and down the country, the cost of assembling and radiating supplies is greatly reduced, and since producers and consumers remain in close contact with one another, there is no need for a long chain of middlemen. In dairying, for example, there is an obvious advantage in having villages and small towns supplied with milk straight from the farm. But if the areas of production and consumption lie some distance apart, so that there is a bottle-neck through which supplies must flow on their way to market, it will be easy for large firms, located at the bottle-necks, to drive small firms, situated on each side, out of business. This point can be illustrated from the livestock industry. In countries and regions which export a large proportion of their meat (e.g., Argentine, New Zealand, or Scotland), the points of export are bottle-necks through which the meat must pass. Thus the cattle can be assembled for slaughter in large meat-packing establishments. In a country like England, however, where both the market and the cattle are spread all over the country, the slaughter-houses are much smaller in size, since the cost of assembling the cattle for slaughter and radiating the meat to local butchers from a centralised slaughterhouse would be very high.[1] In the United States the importance of the export market has declined, and with it the importance of Chicago and the large packers. The expansion of the home market has led to a decentralisation of the industry in smaller plants set up in the leading cattle districts.

Thus the obstacle of distance is strengthened by an increased overlapping of the areas of production and consumption. It is strengthened also by a rise in transport costs and by a reduction in the density of the market or of sources of supply. The trend, however, is generally towards *lower* transport costs and *increased* density. We find bread-making, for example, concentrated increasingly in bakeries in the towns as a result of improved methods of distribution. An area whose bread used to be made in bakeries scattered through villages and suburbs is now supplied by a few large bakers. A further example is provided by the brewing industry, where the substitution of canned for bottled beer, by reducing transport costs, operates in favour of the large firm. Similarly a fall in coal freights, or in the cost of transmitting electricity, may lead to the erection of larger power stations. An increase in the consumption of electricity within a given area (i.e., an increase in the density of demand) will have the same effect.

[1] Nevertheless there is probably far too little centralised slaughtering in England. The sentence in the text is not intended as a justification of the present system of livestock marketing.

Distance, in other words, provides the small firm with a sheltered zone which other firms can with difficulty penetrate. Each small firm, supplying a local market, is partially insulated from competition, but if it tries to push its way into more distant markets, it meets with keener and keener competition round the fringe of invasion.[1] This market resistance to expansion checks the growth of firms to their optimum size. The check will be greater the higher the cost of transport, the more scattered the market and the sources of raw materials, and the more they overlap.

(b) **Psychological.**—This line of reasoning can be extended from geographical to any other limitation of the market. Each firm, as we shall see, is normally marketing a product which differs slightly from the products of its competitors. In the shoe trade, for example, no two firms cater for exactly the same shape of feet, specialise in exactly the same range of sizes, use leather of exactly the same quality, or market their shoes in exactly the same pattern and style, with exactly the same guarantees and advertisements. Each firm has a clientèle of its own, just like the bakers in adjoining villages. The market is broken up, not only by distance, but by the tastes, habits, and prejudices of consumers. Miss A insists on clogs, Miss B on brogues, so they buy from different firms. Mr. Broadfoot finds comfort in the shoes of Messrs. C; Mr. Narrowfoot will buy from no one but Messrs. D. One firm supplies Lady Snooty, another Mrs. Bourgeois, a third John Hiker, a fourth baby Jones, and so on. The market for each firm is limited by the requirements of the customers who are within its "sphere of influence." There is thus a market resistance to expansion which can be overcome only by finding new buyers or by invading the market of other firms. This puts the firm to the expense either of special concessions to its customers, or of an advertising campaign, or of the manufacture of new brands of product. It has to attract new customers by reducing its price, or by selling a better quality of product, or by persuading people that its products are "superior," or by doing something equally costly. These costs of growth, like high transport costs in a scattered market, halt the expansion of the firm below the optimum. They shelter the small firm by penalising invasion of its market. The more varied the attachments which consumers form

[1] For example, a steel company operating a comparatively small plant in Texas had a virtual monopoly of the semi-fabricated steel business in the surrounding territory, being protected by high freight rates from competition from steel-makers in other states. When the size of the plant was doubled in 1926 it became necessary to extend sales into more distant markets, previously supplied by other steelworks. The cost of delivery was greater, while selling prices were held down by competition. In spite, therefore, of an improvement in the efficiency of the plant, the margin of profit disappeared and the company failed at the beginning of 1929 (A. S. Dewing: "Financial Policy of Corporations," page 724n.).

(the less the density of demand), and the less responsive they are to efforts to change their attachments (e.g., by advertisement, or by concessions in service, quality, or price), the more difficult will it be to dislodge small firms from their hold on the market.[1] It is only when consumers are indifferent from whom they buy, and what brands they buy—or when they actually have a preference for dealings with large firms—that the market ceases to be a major obstacle to expansion.

The market obstacle can, however, be circumvented in two ways: by setting up branch factories, or by manufacturing a wide range of products or brands of the same product. The setting up of branch factories is a way of circumventing the geographical limitation of the market; the manufacture of a wide range of products is a way of circumventing the psychological limitation. Neither of these "dodges" is altogether satisfactory.

Branch factories under a central management can supply the adjacent territory without incurring the high transport costs and other marketing expenses of a single large establishment. They can specialise in the articles, brands, and sizes required locally, and supply orders promptly from a stock which is small in comparison with the stock carried by independent units. Each branch can participate in the technical improvements made by the others, and can detect and eliminate waste by comparing its accounts regularly with theirs. Risks are spread, since a period of bad trade in one district can be set against a period of good trade in another, while if production is held up at one plant by fire or strike, customers can still be supplied from other plants. These and similar advantages make for the combination of scattered plants in a single firm when the individual plants are prevented from growing by market limitations.

Distance, however, sets a limit to the combination of plants just as it does to the size of the individual plant. The more scattered the units and the poorer the system of communication between them, the more insuperable are the difficulties of efficient supervision by the central office. If local circumstances differ greatly, responsible and expensive district managers must be appointed. Each plant must be given more latitude, and this makes it all the harder for the central office to keep them in step with one another, and frame a suitable policy. In short, branch factories sooner or later bump up against the managerial obstacle to expansion.

The same obstacle, as we have seen, checks the multiplication of products, patterns, styles, etc., within a single plant. It is only through the sacrifice of great economies in management (and also in technique) that firms can overcome the market obstacle to expansion

[1] A familiar example is retailing, where small shops have an intimate knowledge of the requirements of their customers, and can bind them by the skilful use of credit even when large firms are charging much lower prices.

by extending their range of product.[1] The work of management becomes more complicated, and at the same time the technique appropriate to the large-scale manufacture of a single product or style of product has to be abandoned. More than this, multiplication of products very often fails in its object. For if every firm adopts the same expedient and tries to supply "the whole gamut of varieties" sold by its competitors, it will end by producing everything on a small scale and selling in a market which has shrunk because of the universal rise in costs of production. The public will be offered a wide variety of styles, but it will be denied the benefits of large-scale production, for each firm will manufacture much the same varieties as its competitors, instead of specialising on a limited range of styles and turning them out on a scale approaching the optimum. Sometimes the optimum scale of production is very small, and the economy of concentrating on a few special lines is confined to the time saved in resetting machinery. But more often the failure of firms to specialise more narrowly causes waste from the social point of view, while market resistances to growth simply increase *pari passu* with the variety of output of the typical firm.

It is impossible, for example, to doubt the wastefulness of selling (in the United States) 10,000 brands of wheat flour, 4,500 brands of canned corn, 1,000 brands each of canned peaches and salmon, 500 brands of mustard and 300 brands of pineapple.[2] In Britain, articles in common use--e.g., electric switches, paper bags, envelopes, labels, and string--are sold in an equally bewildering array of brands, patterns, sizes, and colours. A single firm, manufacturing one article in a dozen different sizes, shapes, and colours will turn out no less than 1,728 varieties. If there are twenty other firms, each using slightly different specifications, the number of varieties may be over 30,000. If, finally, each of these firms is simultaneously turning out some ten or a dozen quite different products, the waste and confusion can readily be imagined. Nor is this an impossibly extravagant picture. The engineering trades, for example, include a "very large number of relatively small firms . . . each with a separate organisation, separate establishment charges, separate buying

[1] How great these economies can be may be illustrated from the experience of the Jantzen Knitting Mills which originally turned out sweaters, coats, caps, stockings and many other varieties of "knit wear." Some of these products were eliminated until finally in 1926 it was decided to concentrate entirely on swimming suits of only one quality and to make these on a large scale. "Prior to the simplification programme, an operator sewing a special seam turned out nine seams per hour. Under simplification, by means of special machines and specialisation on this one operation, she produced with no greater effort 45 of these seams per hour. . . . Supervision was made simpler, with the result that foremen and superintendents found time for more work of a constructive nature." Robbins and Holts: "Industrial Management," page 58, quoted by P. S. Florence,: "Logic of Industrial Organisation," page 28.

[2] S. H. Slichter, *op. cit.*, page 553.

and selling arrangements, and each producing a multiplicity of articles. Some of them seemed to take a special pride in the number of things which they turned out. . . . The result of many firms being employed upon producing a large number of articles in common use is . . . confusion in the types of articles produced, so that no two manufacturers seem intentionally to produce the same article. Each one claims some special merit of his own."[1]

Apart altogether from the social waste involved, the production of a whole range of styles, while it may be forced—or appear to be forced—on each manufacturer by the action of his competitors, does not free a firm from the limitations of its market. The pressure to expand may be diffused over a wider area; but so, too, is the resistance to expansion. If one firm markets a new variety of product, other firms tend to follow suit. Even if they do not, they can threaten retaliation in other ways. Suppose, for example, that the new variety sells well. Then unless new layers of demand are tapped, sales will be mainly at the expense of the old varieties, and competing firms will find themselves in difficulties. They must either reconcile themselves to reduced sales or fight to retain their " share " of the market. Very likely, they will fight—for instance, by cutting prices, or by spending more on advertising. The firm which took the initiative will then lose some of its original gains. So long as the threat of retaliation has to be faced, therefore, the marketing of new varieties of product is not greatly successful in overcoming market resistances to growth.

There is a market resistance to expansion on the side of supply as well as on the side of demand. A firm—or more frequently an industry—may find it difficult to obtain adequate supplies of one or more of the factors of production—adequate, that is, for optimum production. Labour, land, materials, or money-capital may be scarce, so that larger supplies can be obtained only at increasingly higher rates of pay. A rise in output will then bring about a rise in costs. Rather than incur such additional cost, firms may be content to remain smaller than the optimum.

### 3. Financial Obstacles.

An important reason for the survival of small firms is the difficulty of procuring sufficient capital. A small firm seeking to expand has generally to finance extensions of plant out of profits, or out of the personal savings of its owners and their friends. It has not a sufficient reputation with the investing public to raise capital through the medium of the Stock Exchange, and does not make large enough profits to permit of rapid expansion. The rise of the joint-stock system, however, has done much to remove the financial obstacle to growth—particularly for firms which are already fairly large. How it has done this, and how far it has been successful, will be discussed in detail in the next chapter.

[1] Quoted from the Report of the Board of Trade Engineering Trades Committee of 1916-17 by P. S. Florence, *op. cit.*, page 27.

There remains one group of forces which we have lumped together under the heading of " desire for competitive strength." These forces occasionally limit the growth of firms,[1] but more usually encourage firms to combine so as to increase their bargaining power. The hope of establishing a monopoly, for example, may push firms into expansion beyond the point at which efficiency is at its optimum. This group of forces will be left over for later analysis.[2]

## CHAPTER 8

## THE FINANCE OF LARGE-SCALE PRODUCTION

### 1. The Need.

BEFORE the industrial revolution, the representative firm was tiny by modern standards, and was owned and managed by one man or by a partnership. The chief industries were carried on on a small scale in the fields or in the cottages, rather than on a large scale in factory and mine. It was only in foreign trade that there was much scope for the large firm. It was in foreign trade, therefore, that the need for large-scale borrowing first made itself felt; and it was in foreign trade that modern methods of finance (e.g., through the joint stock company) were first evolved.

The channels into which savings could flow, or through which they could be borrowed, were narrow. The landowners and the merchants had almost a monopoly of capital and used it on their land and property or in their businesses without much recourse to borrowing. There was little scope for expansion and little incentive, therefore, to accumulate capital or to borrow it. Improvements to property, social display, and mere extravagance swallowed up what the capitalists of later centuries, putting money before magnificence, would have sunk in stocks and shares, or factory buildings and plant. It was not until later, when the agricultural revolution began to create opportunities for investment, that thrift acquired its attraction and its virtue. Nevertheless, even in those days, there were a few spectacular investments which strained the fortunes of the wealthiest landowners. At the beginning of the seventeenth century the Earl of Bedford spent over £100,000 on the draining of the fens, and in the following century the Duke of Bridgewater managed to raise more than twice as much for his canals by cutting down his personal expenses to £400 a year and borrowing from his own tenants to pay wages.

The inventions of the eighteenth and nineteenth centuries, and above all the coming of the railway, increased enormously the scale to which

[1] For instance, in the purchase of raw materials, a firm may wish to avoid buying too large a proportion of the available supply.

[2] See Chapter 9.

a business could profitably grow. But large-scale production without borrowed capital was impossible. The immense sums required exceeded the fortune of even the richest capitalist. Nor was the alternative method of raising capital through partnerships at all suitable. Partnerships are possible only between persons who have complete confidence in one another, since agreements made by any one partner are binding on the partnership. Each partner, moreover, is liable without limit for the entire debts of the partnership, however small the stake which he has in it. If one partner wishes to withdraw, or dies, it will be necessary for the surviving partners either to buy him out—perhaps at great trouble to themselves—or find a purchaser who is acceptable to everyone. These difficulties—the moral and financial risks of partnership and the danger of unexpected withdrawals—multiply rapidly as the number of partners, or the capital which they have to put up, increases. It is impracticable either to raise small sums from a large number of partners, or large sums from a small number. It is necessary, therefore, to find some device by which persons can provide the capital required by a large business without running the risks to which partners are exposed. The limited joint-stock company is just such a device.

## 2. The Solution.

A joint-stock company, or business corporation as it is called in America, is a body corporate with a common seal, carrying on business under the management of a board of directors and owned by a group of shareholders. The company must conform to the requirements laid down by successive Companies Acts, and must file a copy of its memorandum and articles of association with the Registrar of Joint-Stock Companies. After registration, the company enjoys certain privileges which enable it to overcome the legal obstacles to growth that beset the one-man firm and the partnership. First, the moral risk is removed. A shareholder, unlike a partner, cannot bind the company by his acts. The decisions which bind the company are taken by his representatives, the directors, or by the officials appointed by them. So long, therefore, as shareholders have confidence in their directors, they have little need of confidence in one another. Partners require to be in daily contact with one another. But shareholders can sleep soundly in complete ignorance of the identity, and hence also of the honesty, ability and intentions of their fellow-shareholders. Secondly, in a limited company, the financial risk is greatly reduced. Before 1855, when limited liability was introduced, shareholders were regarded by the law as partners, pledging the whole of their property against any debts incurred by the company. If the company failed, therefore, the shareholders might be ruined.[1] In 1855, however, it became legal to limit liability

[1] The failure of the City of Glasgow Bank in 1878, for example, ruined thousands of small shareholders, who had to pay £2,750 in cash for each £100 of their shares.

to the nominal value of the shares, so that, as a rule, the maximum loss which a shareholder stands to make is represented by the capital which he has actually invested.[1] Finally, withdrawals cease to be troublesome. Shareholders are at liberty to sell out whenever they choose, and can do so without disturbance to the work of the company. There is an organised market (the Stock Exchange) in which shares are readily negotiable, so that there is no difficulty in finding a buyer.

The joint-stock company makes it possible to tap the savings of a large number of people, without requiring them to take part in the management of the company. These people provide the capital of the company and receive in return stocks, shares, debentures, etc., entitling them to an income in the form of dividends on their stocks and shares or interest on their debentures. The terms on which the capital is supplied differ for different classes of shareholder. There are differences in income-rights, in the risks which shareholders run, and in the powers of control which they can exercise. The least definite rights to participate in the profits of the company, the greatest risks, and the greatest powers of control are those attaching to ordinary shares. At the other extreme are debentures, which are loans to the company, not shares in it. They carry a fixed rate of interest, comparatively little risk, and no power whatever over the company so long as interest is paid regularly and the rights of debenture-holders are not threatened. Between these limits, as a kind of compromise, stand the preference shares.

**(a) Ordinary Shares.**—In theory, control of a company normally rests with the ordinary shareholders. They elect—or perhaps one should say approve—the directors, and can, by exercising their voting rights, supersede the retiring directors until, after a period of years, a completely new board has been appointed. The ordinary shareholders rank last in their claims on the property of the company, and participate in the profits only after other shareholders have received their share. Thus they bear the heaviest risks.

**(b) Preference Shares.**—Normally, preference shareholders have voting powers only when it is proposed to alter their rights, or to wind-up the company, or when their dividends are in arrears. They are paid, not a fluctuating dividend, but a fixed rate of, say, 6 or 7 per cent. This dividend is not exceeded when the company is making large profits, however high the dividend earned by ordinary shareholders. It may not be paid at all, if the company is in difficulties and failing to earn a large enough profit. The dividends paid to preference shareholders necessarily vary less than the dividends

[1] Some Companies, however, issue shares without calling up their full value. This provides a reserve liability on which the company can draw if threatened with bankruptcy. If, for instance, a shareholder is asked to subscribe only 10/- for a £1 share at the time when a new company is floated, he is liable for the remaining 10/-.

on ordinary shares, and the risks run are correspondingly less. They are less also because preference shares usually rank ahead of ordinary shares if the company fails and its assets have to be liquidated. This difference in risk corresponds to some extent with the difference in voting power and control between the two types of share.

(c) **Other Types of Share.**—This comparatively straightforward division into variable and fixed dividend shares, voting and non-voting shares, is complicated by sub-divisions within each group. There are for example, many different varieties of preference share. There are participating preference shares, carrying the right to an increase in dividend if dividends on ordinary shares rise beyond some stipulated figure. Then there are cumulative and non-cumulative preference shares. On the first of these, but not on the second, dividends passed in one year must be made up in later years, and so long as there are any arrears, ordinary shares cannot receive any return. It may happen, however, that the company has to undergo financial reorganisation in order to avoid winding up altogether. Arrears even on cumulative preference shares may then have to be sacrificed. Finally, some preference shares are redeemable. This is true only of shares issued since the 1929 Companies Act, so that most preference shares are still irredeemable (i.e., they cannot be paid off).

The sub-divisions of ordinary shares are still more complicated. There are preferred ordinary shares and deferred ordinary shares the one ranking ahead of the other until dividends reach some agreed limit. There are non-voting ordinary shares designed to leave control of the company to the holders of special voting shares. And there are shares carrying a voting power out of proportion to the capital invested. These shares, too, concentrate power in the hands of persons who may own no more than a small fraction of the company's assets. A device which is even more effective in divorcing ownership and control is the voting trust, which is common in the United States. Shareholders surrender their voting power in the company to a group of trustees (usually the management) in return for voting trust certificates entitling the holders to a share in such dividend payments as the directors choose to make. The fruits of ownership are thus enjoyed by men who, even as a group, have lost the power to control their own enterprise.

In Great Britain all shares must have a nominal or " par " value, which is normally in rough correspondence with the sum invested by the original purchaser of the share. The market value of shares may differ widely from their " par " value if the company has been either extremely successful or extremely unsuccessful. The holder of a £1 share may realise only 6d. or as much as £10 when he sells it, depending upon the profits and prospects of the company. In America the issue of ordinary shares of " no par value " (i.e., titles to a certain proportion of the distributed profits of the company) has been legal for many years in most States, and shares with a nominal value are now rarely issued.

(d) **Debentures.**—The capital of a company is not provided solely by its shareholders. Borrowings are made through the sale of bonds or debentures, and this borrowed capital goes to supplement the share capital. Debenture-holders are not, like shareholders, proprietors of the company.[1] They are creditors with no voice in the management and policy so long as they receive interest on their bonds. The debentures which they hold are generally repayable after a stated term of years at a fixed price, and form a first charge on the assets of the company. As security for the debt, the debenture-holder has not only all the assets in the possession of the company at the date of the debenture, but also all the assets coming into its possession until the loan is repaid. When the debenture stock is issued, trustees are appointed to act on behalf of the holders, and these trustees can, if default on interest or principal takes place, secure the appointment of a receiver, who is authorised to take possession of the property of the company and to manage or sell it until the loan is repaid.

### Investment and Risk

The joint-stock system, then, increases the supply of capital at the disposal of a single business, and so removes one of the chief obstacles in the way of large-scale production. But large-scale production requires that capital shall be available not only in large amounts, but for long periods, and for risky undertakings. Large businesses almost invariably use much fixed capital. This capital is, so to speak, highly specialised—that is, sunk in forms from which it will not be recovered for many years (ploughs as compared with seed), or in forms with a very limited number of uses (battleships as compared with crude steel). The longer the life of the fixed assets (the buildings and plant), the longer the period for which investors must part with their capital; and the more highly specialised are the fixed assets (either because they are durable or because they are not adaptable) the greater will be the risks which investors run. It may prove impossible to make regular use (or indeed *any* use) of the fixed capital over the period of its life; or it may yield a return lower than was expected; or it may yield a return which, while up to expectations, is less than might have been obtained by holding the capital in a more liquid, less specialised form and taking advantage of a change in circumstances favourable to some other line of investment. Whatever the *result* of sinking capital in fixed assets, the uncertainty is great; and the large firm, with the bulk of its assets fixed and specialised, has given an unusually high quota of hostages to fortune and has given them irrevocably for an unusually long time.

---

[1] Nevertheless their position differs very little from that of holders of first-class, redeemable, non-cumulative, non-participating preference shares. Sometimes, too, convertible debentures are issued, i.e., debentures convertible into shares at the option of the holders. In such a situation the distinction between dividends and interest payments is a very fine one.

Add to this the risks of producing in advance of demand. A small business can often make to order without risk of loss. But a large business supplies a market remote in time and space—a market whose requirements cannot be accurately foreseen. The wants of this market lie in the future and at a distance, but production must be undertaken here and now. The whole organisation of the business, and not just its fixed capital, is built and operated in ignorance and uncertainty of the future demand. The owners must adventure their capital over the period of production and bear the risk of loss through error or misfortune.

There are thus two difficulties which the joint-stock system has to overcome. The first arises from the desire of the investor for liquidity. He wishes to be able to recover his capital quickly and without loss; but it has been sunk in durable plant, and can be recovered (in depreciation allowances) only over the life of the plant. How can the needs of the investor be met? The answer is that although the asset is illiquid, the investor's shares are not. There is a market in which they can be sold at any time (but not necessarily at their purchase price)—the Stock Exchange. The function of the Stock Exchange is to make stocks and shares easily marketable. If there were no Stock Exchange, shareholders who wished to sell out would be in the same position as holders of a mortgage, or of shares in the village gasworks. They would be put to the trouble of advertising, or of negotiations with a lawyer, before they were able to find a purchaser. Even when a purchaser had been found, he might offer no more than a knock-down price. So long as the market for shares was so poorly organised, people would be reluctant to lock up their money in stocks and shares. The Stock Exchange, by improving the marketability of shares, correspondingly encourages investment in them. Since dealings are regular, the investor can sell at a moment's notice at a price which is uniform throughout the market. Since dealings take place over a wide area on local exchanges connected with one another by telephone and cable, very large blocks of shares can be sold without greatly depressing their price. Thus the market price of shares, in normal circumstances,[1] is steadier as well as more uniform when dealings are organised than when they are not.

The second difficulty to be overcome is the disinclination of investors to expose their capital to risk of loss.[2] This disinclination is reduced by a fourfold specialisation in risk-bearing.

First, there is a specialisation between directors and shareholders. Many owners of capital show little aptitude or inclination for the work of managing a business, while many capable business men

[1] This qualification is intended to cover speculative manœuvres rather than industrial fluctuations between prosperity and depression. But even allowing for these manœuvres the price of traded shares is likely to be steadier than the price of non-traded ones.

[2] No judgment is passed above upon whether the Stock Exchange performs its functions adequately or not. On this, see pages 99-101.

have little or no capital. The joint-stock company, in divorcing ownership from management, simultaneously unites management with capital. Capital and business ability are brought together, and the resources subscribed by shareholders are put at the disposal of directors who are presumed to possess good judgment and organising ability. The union of capital and business ability renders a double service. On the one hand, more money is saved because acceptable channels of investment have been created; and, on the other hand, the limits within which business men are able to exercise their judgment are expanded. There is more capital available, and more of it is loaned to, or invested in, large businesses.

Secondly, the practice of issuing different kinds of shares caters for the different dispositions to take risks of different classes of shareholders. There is a specialisation in risk-bearing, not only between directors and shareholders, but also between debenture-holders and shareholders, and between preference, ordinary and deferred ordinary shareholders. This grading of bonds and shares according to the risks which the owners run allows the enterprise of investors to be harnessed more effectively in support of risky undertakings. The supply of risk-bearing, as it is sometimes put, is increased.

Thirdly, the participation of a large number of shareholders in each company limits and diffuses the risks which they run. A man with a total capital of £10,000 may scruple to sink three-quarters of it in a partnership with a dozen others. He may hesitate even more to invest the same sum in a company over which he has no control. But if, instead of being one of twelve, he is one of a hundred or two, and is asked to put up, not £7,500, but a few hundred pounds only he will have much less hesitation. If he chooses, he can invest in a large corporation like the American Telephone and Telegraph Company with over half a million investors—more investors, indeed; than employees.[1] He can then invest the rest of his capital in other businesses and spread his risks on a large number of shareholdings; His risks will be still further reduced by the limitation of his liability in each investment to the value of his shares. The shareholder, therefore, can limit his risks in three ways: by limiting his stake in each investment, by investing in a number of different companies, and by enjoying the protection of limited liability. This limitation of risk will once again make it easier for large or risky undertakings to raise capital.

Finally, there are various financial intermediaries which are enabled by the joint-stock system to borrow small amounts on good security and lend large amounts where the return is comparatively uncertain. The deposits of the joint-stock banks, for example, consist very largely of small sums, withdrawable at short notice. The assets,

[1] Such a company literally owes its existence to the joint-stock system, since not even the biggest millionaire owns more than about one-tenth of its £850 m. capital.

on the other hand, include advances, for periods up to six months or a year, of very large amounts, and blocks of Government bonds, not redeemable for several years, running into hundreds of millions of pounds. The banks' depositors lend their money in complete safety and without trouble; the deposits are lent by the banks after a great deal of trouble to discover which firms can offer satisfactory security, and not all of the loans are recovered in full.

Insurance companies and building societies perform similar functions. They receive a large proportion of the current savings of the public, the insurance companies in premiums on life assurance, and the building societies in repayments of principal on mortgages, or in deposits or from shareholders. These savings, when pooled together in this way, can be lent for longer periods, or in larger amounts, or to riskier undertakings, than when they were in the hands of individual savers.

Another important agency in achieving these results is the investment trust. An investment trust aims at spreading the risks of its shareholders by distributing its assets over a large number of different shares—either shares in different companies or in the same industry and area, or, more commonly, in different companies scattered throughout the whole range of industry and over the whole area of the globe. The purchase of shares in an investment trust, therefore, allows a small capitalist to take an interest in a large number of enterprises exposed to a variety of risks, instead of linking his fortunes too closely with the one or two companies only in which he could otherwise afford to take shares.

These investment intermediaries, with the exception of the banks, either did not exist or were of negligible importance less than a hundred years ago. To-day, the combined assets of the joint-stock banks, savings banks, building societies, insurance companies, and investment trusts amount to nearly £6,000 millions. The holdings of stocks and shares by the last two groups probably amount to a tenth of the total held in Great Britain, and this proportion is increasing rapidly.[1] Moreover, the savings made through investment intermediaries, including the huge sums put to reserve every year by the joint-stock companies themselves, are more than sufficient to account for the total annual savings of this country.[2]

[1] A. T. K. Grant: "A Study of the Capital Market in Post-war Britain," page 198.

[2] In boom years, undistributed profits amount to about £300 m. In 1934—a poor year—they are estimated to have reached £156 m.; repayments of principal to building societies came to £72 m.; and savings through insurance companies or by deposit with the Savings Banks, purchase of Savings Certificates, and so on, amounted to about £100 m. In all, therefore, we have a total of £328 m., which is more than £100 m. in excess of net savings in 1934. (For these estimates, see Colin Clark: "National Income and Outlay," pages 185-193).

**Diffusion of Ownership.**

Stock Exchange securities are owned by three broad groups: by private persons, by investment intermediaries, and by large companies as part of their reserve. Of these groups, the first is still far and away the most important. The banks (including the savings banks), the insurance companies, and the investment trusts hold between them perhaps £3,000 millions in Stock Exchange securities. Fifteen large companies, including the four main line railways, held £125 millions in securities in 1935.[1] But the total value of Stock Exchange securities held in Great Britain is at least £18,000 millions.[2]

If we exclude Government bonds and take securities in joint-stock companies only, the preponderance of private investors is increased rather than diminished. Company reserves are held mainly in the form of Government securities, which we are now leaving out of account. The security holdings of the banks must be almost entirely omitted for the same reason. It is safe to say that for every £100 of securities in joint-stock companies which are owned by financial institutions or by companies engaged in production and distribution, at least £1,000 of securities are in the ownership of private persons.

These private stockholders are a numerous class. In the United States it has been estimated that between four and seven million persons own stock.[3] In Britain stockholders can hardly number less than two millions.[4] Not only is the ownership of Stock Exchange securities widely diffused, but so too is the ownership of any one company. The largest British industrial company, Imperial Chemical Industries Ltd., has well over 100,000 shareholders. The railways, with an even larger capital, reach still higher totals. The stake of the average shareholder in any one company is thus comparatively small. In a sample analysis of ten of the largest British companies, made in 1932, less than an eighth of the ordinary shareholders held £500 of nominal capital or over.[5] The average holding of ordinary

[1] Grant, *op. cit.*, pages 196-7. The companies (other than the railways) were: Courtauld's, Shell Transport, Imperial Tobacco, P. & O., I.C.I., Vicker's, Anglo-Iranian, Distillers, Guinness, J. & P. Coats, and Morris Motors. The security holdings of these companies came to over a third of their nominal issued capital.

[2] The value of securities entered on the Stock Exchange Official List is over £18,000 m.; a similar total is reached by adding together the capital of companies registered in Britain, the debt of the local authorities and of the central government, and our loans to governments and municipalities abroad.

[3] Berle and Means, "The Modern Corporation and Private Property," Appendix K.

[4] The number of persons earning incomes of over £250 per annum is just over two millions. For everyone above that level of income who owns no shares there must be at least one person below.

[5] H. Parkinson: "Scientific Investment," page 8. The shares of many of the companies were selling above their nominal value, however, so that, judged by market value, the proportion of investors with a holding worth £500 or more may have been higher than an eighth.

or preference shares in a large company has generally a nominal value of about £300.[1]

**Absentee Ownership.**

This mob of owners—scattered, ignorant, apathetic—has rarely any interest in its property beyond the squeezing from it of the maximum possible return. Practically the only rights of ownership which it can exercise are those of attending and voting at the general meetings of the company. The responsibilities of ownership it either does not recognise or is impotent to assume. A small factory owner, the sole proprietor of his business, knows that he has other duties—to his men, to his customers, to the public—than merely to keep within the law; and other rights—of command, of appointment and discharge, of organisation and control—than the right to make and enjoy such profits as he can. He is the dictator in a small kingdom, with power, if he chooses, to be a benevolent dictator. But what of the absentee dictators who rule (nominally at least) the vast empire of a large joint-stock company They may know nothing of how their company is run, or what kind of life men lead under their despotism. If they know and disapprove, and are willing, even at the sacrifice of profit, that some improvement should be made, how can they force their views on the management? They have thousands of other stockholders to convince—strangers who, as like as not, will be sceptical or indifferent, and who can be reached only at great expense by letters which may not even be read. Shareholders find it hard enough to override the policy of the management when they believe their personal interests to be threatened. How much harder must it be for them to obtain support for a policy from which no immediate benefit in increased profits is expected—through insistence, for example, on decent working conditions?[2]

[1] This low average is often cited to prove that capital is fairly equally distributed. But one might as well cite the low consumption of bread by the rich to prove that they are not really rich at all. The average investment is low, because capitalists do not like to put all their eggs in one rickety basket. The vast majority of workers put their solitary egg in the very sturdy basket of the Post Office Savings Bank or of a Building Society, and have never owned shares in their life. The two million persons who (at a guess) own shares in Britain form no more than a seventh of the adult male population. Few of these two millions own more than a few hundred pounds worth of shares. In fact, of all forms of capital, stocks and shares are *most* monopolised by the wealthy. In the United States, three-quarters of the dividends paid by corporations went in 1929 to 600,000 persons (in a population of 130 millions). There is no reason to suppose that the distribution of ownership is more equal in Britain.

[2] This must not be taken to imply that the pay and working conditions in large limited companies are in fact worse than in small private firms. The reverse of this is true. The large company cannot risk the goodwill of its customers by exploiting its workers, while small firms, not working in the full glare of publicity, very often can. The large firm, too, must offer more attractive terms to its workers just because the men prefer working for a boss

**Speculation.**

The owners of a company are not only absentees. They do not even remain the same absentees. They are a shifting body of men, like hotel guests, in to-day and out to-morrow—or sooner. They buy and sell again, not because death and necessity part them from their capital, but because they see a chance of a speculative profit. Very often, they have no interest in the ultimate destiny of their company, but are satisfied if the price of its shares over a period of days or weeks moves up or down as they anticipated. At this stage, they have ceased to be property owners with a stake in the company and entitled to the same rights as other property-owners, and have become speculators pure and simple.

Now *speculation is inseparable from all investment*; every owner of capital must "take a chance" in whatever form he holds his capital. Risk is universal. Speculation, too, may be of immense advantage to society. Every pioneer of a new process "speculates" his capital on its success. Every storer of wheat who buys the surplus crop of a bumper year speculates on a short supply in the succeeding years. Thus skilful speculation can work improvement in our use of scarce resources, or reduce the risks (e.g., of famine) to which their scarcity exposes us. The speculator is not just a gambler, but a specialist in risk-bearing who, if he bears necessary risks wisely, saves us from those who are more imprudent, or less capable (through ignorance, or lack of capital, or excess of caution) of bearing the risks which they run. But the risks must be unavoidable, not the product of bad organisation or a desire to encourage gambling; and the speculators must be well-informed and sound in judgment, not acting on "inside" information, which others do not share, nor seeking to spread rumours and false impressions, nor backed by such weight of capital that they can corner the market and reap monopoly profits.

How far are these conditions satisfied on the Stock Exchange? That the Stock Exchange is intended to reduce unavoidable risks we have already seen.[1] It makes it less risky to lock up money for long periods, and thus increases the flow of capital into productive investment. But there are also risks on the Stock Exchange which are, so to speak, fictitious. The operations of the "bears," who speculate for a fall in prices, are designed to alarm the public and create a sense of risk of which the "bears" can take advantage; the operations of the "bulls," who speculate for a rise in prices, are designed to make risks appear less than they really are, with equal advantage to the speculators. Thus even if professional speculators,

"in the flesh" to working for absentee bosses. Dissatisfaction expressing itself in slackness, poor work, or a high labour turnover can be very costly in a large business, and the cheapest way to avoid dissatisfaction is to offer good working conditions. Thus self-interest may accomplish what a sense of the responsibilities of ownership is powerless to do.

[1] See above, page 94.

or those whose interest is predominantly speculative, were exceptionally judicious and capable of keeping market prices in correspondence with the real risks, they are not free to do so. It does not pay to steady market prices when by unsteadying them there is a better chance of profit. The real risks become overlaid by the fictitious risks, and wisdom dictates a forecast, not of what prices should be, but of what, given the psychology of the market, prices will be. We reach a stage at which the value of shares reflects, not a genuine expectation of their yield over a long term of years, but the views that average opinion expects average opinion to take within the next few weeks, or days, or even hours.[1]

Nor do professional speculators rely only on their superior wisdom. The use of rumour and of "inside" information is frequent. The information may come from the directors and officers of the company in whose shares speculation is taking place, or from politicians.[2] Or the speculators may themselves be directors and officers or politicians. In the United States—where graft (and, it should be added, exposure of graft) is on a bigger scale than in Britain—it was once not uncommon for directors to abuse their privileged position by withholding important information from the public while they speculated in the shares of their company. In Britain, commercial morality is not on such a high level that such "privileged" speculation is unknown.[3]

Speculation with a view to cornering the market, although it attracts great attention,[4] is rare and generally unsuccessful. In the commodity markets, the capital required is enormous, and new supplies may be forthcoming to overstrain the resources of the speculators.[5] On the Stock Exchange, cornering stock simply means buying out other shareholders—and so is not very dangerous to

[1] Lord Keynes: "General Theory of Employment, Interest, and Money," page 156. Compare also page 160 n.: "It is said that when Wall Street is active, at least a half of the purchases or sales are entered upon with an intention on the part of the speculator to reverse them *the same day*."

[2] The famous Budget leakage of 1936 may be exceptional. When the Khedive's shares in the Suez Canal were purchased by Disraeli in 1875 there seems to have been a similar leakage (L. H. Jenks: "Migration of British Capital.").

[3] Naturally it is not easy to give examples. It is alleged, however, that before the Marconi Company was bought by the Cables and Wireless Combine at a price which caused its shares to jump from twelve to eighty shillings, purchases were made by financiers with inside information as to the company's affairs. (Margaret Miller and Douglas Campbell: "Financial Democracy," page 44 n. For another example, see also page 117).

[4] For example, the Leiter wheat corner of 1898 and the Bishirgian pepper pool of 1934. Only the first was successful, and its success was due in large measure to the poor crop which followed the corner (that is, the holding back of supplies was, after all, of real advantage).

[5] As in the Bishirgian case.

anyone—or it means putting one speculator in a position to squeeze another who has oversold—and again is not a menace to the public.

It is sometimes suggested that Stock Exchange speculation is pure gambling. But gambling only makes money change hands; speculation makes not only money but also stock change hands, and reacts on prices. When money changes hands, only the gainer and loser (and their dependants) are affected. But when the price of stocks and shares moves up or down, there is a long chain of repercussions on economic life. The investment of capital, the cost of living, the amount of unemployment, are all tied up with stock prices. There is a saying in America that "trade follows the ticker." It is a Wall Street saying, and Wall Street is the land of shibboleths and myths. But for once, Wall Street is right. Trade does follow the ticker. Not, it must be said, in the Wall Street sense that a boom on the Stock Exchange is enough to generate a boom in industry; but in the sense that capital flows into those trades whose shares speculators are prepared to buy, and out of those trades whose shares speculators are pressing to sell. Men's jobs, too, depend upon the course of speculation. Where capital cannot be raised, men cannot find employment, and the ease with which capital can be raised depends not only on the real prospects of return on it, but also on the mood of speculators.

It is this, more than any damage to our morals which the get-rich-quick atmosphere of speculation may inflict, that is the most damning criticism of the Stock Exchange. It diverts men's energies from the tasks of enterprise—of building intelligently for a future which we can only dimly foresee—to speculation for its own sake—trying "to guess better than the crowd how the crowd will behave."[1] Instead of promoting the capital development of a country along socially advantageous lines, it ties it to a market whose chief preoccupation is what its chief preoccupation is likely to be. Capital development becomes "the by-product of a casino."[2] Hence the agitation for a National Investment Board, under the control of the State, to direct the savings of the country into the proper channels.

**Concentration of Power.**

The joint-stock company was originally a democratic organisation. In intention it is a democracy of capitalists, entrusting their surplus funds to a cabinet of their own choosing, and retaining an active interest in the use to which their capital is put. In practice, it has come to be the means by which financial and industrial power is concentrated in the hands of a small number of persons who are only vaguely and on rare occasions responsible to their shareholders. Industrial democracy—even a democracy of capitalists—has receded further and further as political democracy has come nearer and nearer,

[1] Lord Keynes, *op. cit.*, page 157.

[2] *Ibid.*, page 159.

until, with the winning of manhood suffrage, all the important decisions in industry and finance have come to be made by a few score of directors, most of them known to one another.

This concentration of power is not the product of a simultaneous concentration of wealth; capital is not less, but more, equally diffused. Power has been concentrated because the bonds between ownership and control have been loosened. It is not necessary to own the entire capital of a company in order to control it. It is enough to have "a controlling interest." This may be secured through ownership of a bare 51 per cent. of the ordinary shares of the company. If, for example, capital is raised through the issue of debentures, preferred stock, and ordinary shares each to the value of a million pounds, the comparatively modest investment of £500,100 in ordinary shares will carry with it control of the company's £3,000,000. Generally, however, a controlling interest can be purchased much more cheaply. The policy of the company can be dominated by a strong minority, holding perhaps no more than 10 or 20 per cent. of the voting shares, so long as there is no rival minority of equal or greater strength, and so long as the management is not antagonistic. The larger the number of shareholders, and the more diffused the shares of the company amongst them, the harder does it become to take concerted action to dislodge a powerful minority from control.

Many companies have reached such a size that control either by a majority or by a strong minority is impossible. The largest single interest may amount to no more than a fraction of 1 per cent. of the company's capital.[1] In such companies, and in those where there is in fact no large and organised minority group, control normally rests with the management. The directors can remain quietly in power, smothering protests, and procuring their own reappointment by the use of proxies.[2] Shareholders who attack the management can be branded as "agitators" and "self-seekers," and shouted down at company meetings. They are put to the expense of preparing and issuing circulars, and run the risk of depreciating the value of their shares by criticism damaging to the company's reputation. Thus even the powerful investment trusts, with a large holding in the company, rather than fight a policy which they consider unfair and unwise, generally prefer to take the easier course of selling out at a sacrifice price. The meetings of the company where shareholders

[1] In 1929 the largest stockholder of the Pennsylvania Railroad Company held one-third of 1 per cent. of the total shares, and the combined holdings of the twenty largest stockholders fell short of 3 per cent. of the total. (Berle and Means, *op. cit.*, page 85).

[2] Shareholders have the option of voting in person, or refraining from voting, or signing a proxy transferring their vote to a nominee of the management. Most shareholders cannot attend company meetings—if they did they could not be accommodated—and there are generally enough proxies to dispose of the opposition.

are supposed to assert their authority are either deserted,[1] or, when a fruitless struggle for power is on, bear gardens.[2]

There are, in addition, various legal devices by which control can be concentrated. Of these the most striking is known as "pyramiding." One company may obtain a controlling interest in another, which in turn has a controlling interest in a third—and so on through a long chain of companies. The first, and some of the intermediate companies, will generally be holding companies, not engaged in production, but holding the securities of other companies, collecting the dividends on these securities, and in a position to govern the policy of the subsidiary companies. At the base of the pyramid will be a number of large operating companies, while at the apex will be a comparatively small holding company, dominated by a single interest. Complete control over the entire property of the companies grouped together in the pyramid rests with the holding company at the top and with the persons controlling it. The use of debentures, preferred stock, non-voting and weighted shares extends the area of the pyramid. We reach a position in which an investment of less than twenty million dollars secures control over eight Class I American railways with combined assets of over two billion dollars.[3]

Concentration of control has gone furthest in the United States. It has been estimated that, of the 200 largest companies (other than banking companies) in that country, two-thirds are controlled by the management or by legal devices like pyramiding and the issue of voting trust certificates, and that this group of companies owns four-fifths of the combined wealth of the 200 companies. Only 6 per cent. of the wealth is in companies where a majority of the shares are in the hands of a single owner or compact group of owners.[4] For Britain, there are no comparable estimates. It should be remembered that in both countries there are a very large number of small private companies where ownership and control are still combined. But although important numerically, these small companies do not own a very large proportion of the total capital of the country or employ a very large proportion of its workers.

**Division of Interests.**

The policy of any firm reflects a compromise between the various controlling interests. There are divisions of interest between large shareholders and small, between the owners and the management,

[1] For example, at the meeting in 1919 when the Standard Oil Company of New Jersey decided to issue $100 millions of preferred stock, a single shareholder was present. (Slichter, *op. cit.*, page 160).

[2] One such meeting was described in the *Western Mail* as "stormy, full of personal references, and lasting five hours." (Miller and Campbell, *op. cit.*, page 33).

[3] In the Van Sweringen system—which has since collapsed. See Berle and Means, *op. cit.*, page 73.

[4] Berle and Means, *op. cit.*, page 94.

between holders of shares and holders of debentures. How these divisions are reconciled depends upon the legal rights which each group possesses, and on the pressure which they can bring to bear on the company.

The large shareholders, for example, may seek to manipulate prices so as to increase the profits of some other company in which they have a still larger interest. The smaller their holding, the greater the temptation. They may exploit the company by selling materials to it at exorbitant prices or by forming a selling agency which draws handsome commissions for marketing the company's products. They may buy up or build small plants and re-sell them to the company at a profit. They can obtain inside information and use it in Stock Exchange speculation. If they control several companies, they can make their speculations doubly profitable by switching business from one company to another so as to mislead investors. Deliberate financial mismanagement, even to the point of driving the company into bankruptcy, was almost a regular feature of pre-war American railway finance. There is no end to the ways in which those in control can, if they are so minded, exploit those who are not. The more diffused the capital of the company, the more helpless is the small shareholder. In Britain his only weapon —apart from seeking protection through a costly action in Chancery—is that of boycott. If exploitation is too open, no one will take shares in the company; if it is carried too far, profits will be seriously reduced, and again it will be hard to persuade other shareholders to participate.[1]

Conflicts between the owners and the management may arise through exploitation of the same type. The officers of the company may pay themselves high salaries or speculate on inside information. They may conceal losses made by the company so as to prevent a bankruptcy which would lose them their job. They may reinvest profits in other companies so as to increase their own power, and without consulting the interests of their shareholders. Or they may offer higher wages to their workers, or better service to their customers, out of pride in their company—a pride which the shareholders, even if they share it, may not wish to see indulged quite so generously.

Similarly, there is a division of interest between the different classes of investor in the company. The ordinary shareholders will support a riskier policy than is to the advantage of the holders of preference shares and debentures. For whereas these investors, if things turn out well, gain nothing, and if things turn out badly may lose their

[1] An interesting example of exploitation by "insiders" was the organisation of railway equipment companies by the officers and directors of American meat-packing concerns. Railway companies were threatened by these "insiders" with withdrawal of valuable traffic and coerced into buying equipment at inflated prices. The "insiders" were able to use their control over one company to make a profit for another—their own. (A. R. Burns: "The Decline of Competition," page 173).

capital, the ordinary shareholders stand to gain a great deal, and, if they lose, can throw part of the burden of loss on the other groups. If, for example, a company is in danger of bankruptcy, the ordinary shareholders may be willing to run it at a loss, by allowing plant to depreciate without adequate expenditure on upkeep and renewal. By such a policy the ordinary shareholders can lose nothing, and have hopes of tiding over a bad period until there is again a chance of making a good profit. But the debenture-holders see the fixed assets, which are the security for their loan to the company, depreciating steadily. Their interests are threatened, and their trustees may be forced to take action.

**The Growth of Professional Management.**

With the growth in the power of big business has gone a simultaneous growth in the importance of the hired manager. Very often these managers or directors are able to assume control and shake off all but the appearance of responsibility to their shareholders. Any judgment on the value of the joint-stock system, and on its possibilities, must turn to a large extent, therefore, on what we think of these directors. Are they appointed for their efficiency or mainly through nepotism? Are they able and enterprising? Have they evolved professional standards of management that can be commended as in the public interest? Are they, in short, fit men to be entrusted with the immense powers which they wield over the life of the community? Or should the law be invoked to limit these powers and enforce other methods of appointment? Should they, perhaps, be appointed by the State? Or is the whole system of private enterprise, of which they are part, to be completely remodelled in some kind of Socialist State?

Of these questions some only can be answered at this stage. There is no doubt that, as a body, the directors of British companies leave much to be desired. They are drawn—especially those in the larger companies—almost exclusively from the public schools. Of 82 bank directors and 50 railway directors, for example, 53 and 32 respectively attended the 14 most exclusive public schools.[1] Thus they are drawn mainly from one class in the community, and put through one type of educational machinery. The class is not one which has a monopoly of business ability, or which can distinguish with special acuteness between self-interest and public spirit. The education is not one which fits men for business administration by arousing an enthusiasm for hard intellectual work, and by developing powers of abstract reasoning and of rapid assimilation of records and reports. Merely to be well-connected, skilled in the social graces, and trained to regard business as a kind of inferior substitute for cricket is only

[1] R. H. Tawney: "Equality," Appendix I, quoted by P. S. Florence, *op. cit.*, page 252. Another study showed that, in 100 large British companies, 40 per cent. of the directors were titled (including Sir as a title). See Miller and Campbell, *op. cit.*, page 78.

too often a sufficient recommendation for a business appointment.[1] Meanwhile the children of poor parents, who have not the advantage of nepotism, influence and a public school education, are almost automatically disqualified whatever their abilities. They are diverted into the civil service and the teaching profession—in which, by comparison with other countries, there is now as great a superfluity of brains as there is a deficiency in business.

Not only are the avenues of recruitment narrow and twisted; they are also prodigiously long. The directors of British companies are almost all over 45, and some are over 90. In a sample of 804 directors in 100 large British companies only 46 (or 5.7 per cent.) were found to be under 45. On the other hand, 214 (or 26.7 per cent.) were over 70. The average age was 63.[2] When we find the Bank of England retiring its directors at 70, and universities their professors at 65, we may safely disregard the example set in politics by the Gladstones and Lloyd Georges, and conclude that earlier retiral and earlier promotion would put more "ginger" into British industry without any sacrifice in judgment and experience.

## CHAPTER 9

## THE GROWTH OF LARGE-SCALE PRODUCTION

THE fact of growth in industry is beyond question. There is a steady growth in the size of the typical establishment and a steady growth, too, in the number of establishments united in the typical firm. Large-scale production is spreading. But although the tide of growth is sweeping firms into larger and larger units, it does not sweep all firms along at the same rate, and there are undercurrents which carry some firms back into a decline, or wreck them altogether. One firm may grow from a cycle shop into a great motor works inside twenty years, while a flourishing steel-works may be left derelict after five. The fortunes of individual firms are not governed by the general upward trend.

**Motives and Methods.**

A firm may seek to expand from two fairly distinct motives. It may be attracted by the prospect of lower costs of production (resulting from economies of scale), or by the prospect of higher

[1] Given the social inequalities of capitalism, nepotism is not necessarily disgraceful, since a boy brought up in the atmosphere of the family business has great opportunities of acquiring knowledge and contacts which more than compensate for lack of ability.

[2] Miller and Campbell, *op. cit.*, pages 95 and 98. The companies under analysis controlled well over £2,000 millions. The high average is partly to be explained by the gaps created by the last war in the younger age-groups.

prices (resulting from a bargaining advantage or the winning of a position of monopoly). At the same time a firm may seek to expand along either of two paths. It may extend its own plant and fight for a larger share of the market; or it may try to buy out, or combine with, other firms which have their own plant and organisation. Between the two motives to growth and the two methods by which growth takes place there is no necessary connection. It is true that the monopoly motive is generally most prominent when a firm grows by combination, while the economy motive is generally the dominant one in growth by extension. But the difference is one of degree, not of kind. Firms may and do combine in order to reduce their costs; they may and do extend their plant so as to entrench themselves more firmly in a position of monopoly. Growth towards the optimum or towards monopoly may take place along either route.

Whichever method of expansion is adopted, both motives are commonly present. Monopoly gains are always an incentive to expansion; savings in costs can often be made simultaneously, so that the first motive is reinforced by the second. But sometimes expansion may involve the sacrifice of possible economies for the sake of competitive strength or monopoly. The two motives then come into conflict. For example, it may happen that expansion along one route is accompanied by lower costs, and along another route by higher prices. A firm may hesitate between setting up small high-cost branch factories in outlying districts where competitors are at a disadvantage, and building extensions at low cost to its main plant alongside its chief competitors. Or it may have to choose between extending its plant and buying out some of its competitors. Plant extension may improve efficiency greatly, but it will also intensify competition and probably force down prices. Expansion by purchase, on the other hand, while it may be relatively costly and give rise to technical and managerial difficulties, is likely to strengthen the firm's hold on the market and maintain prices or even allow them to be increased. The extra expense of amalgamation must then be weighed against the higher prices and larger receipts which are in prospect. The two motives to expansion, becoming tangled up with alternative methods of expansion, come into conflict with one another.

When such a conflict arises, the social interest is comparatively clear. The monopoly motive is, in general, anti-social; while the economy motive is, in general, in the public interest. So long, therefore, as we can readily distinguish between expansion aiming at monopoly and expansion aiming at economy, we can, by legislating against monopolistic expansion, allow firms to grow in size only when such growth is unmistakably in the public interest. But in practice the two motives to expansion generally reinforce one another rather than conflict with one another, and the public interest, therefore, is by no means obvious. The larger the size of the optimum, the more closely do efficiency and monopoly coincide,

until, when there is room for only a single firm of optimum size, we cannot have one without the other. Efficiency and monopoly become inseparably linked, and we have to lump what we dislike for the sake of what we like. Alternatively, we have to resort to State control or State ownership, rather than leave in private hands powers which may be used against the common good.

The difficulty of disentangling the two motives has increased steadily with the constant enlargement of the optimum by improvements in industrial technique and business administration. So long as the optimum remained small, the distinction between the two motives coincided broadly with the distinction between the two methods of expansion. The growth of independent concerns could be taken as a sign of superior efficiency from which the public would benefit in lower prices. Combination, on the other hand, was a device for getting prices up, not—except as an afterthought—for re-organising production so as to get costs down. It was rarely a means of stepping up the size of the business unit into line with the optimum, but very often a short cut to monopoly. The optimum being small, there was room for several competing firms and monopoly had to be reached at a bound or not at all. Thus it was possible, broadly speaking, to block attempts at monopoly by legislation against combination.

Where the market was limited or scattered, however, there were local monopolies in the growth of which combination had played no part. Of these monopolies the local chemist's shop or the municipal gasworks were typical. These monopolies, free from the " taint " of combination, were regarded as natural and advantageous. It was obviously stupid to have two chemist shops or town gasworks, for the sake of competition, when one would do perfectly well. Competition involved wasteful overlapping of effort and duplication of capital. Now what applies to local monopolies in a limited market applies also to national monopolies when a very large optimum cannot be reached for lack of a market. It is wasteful to build two continuous strip rolling tin-plate mills when one is big enough to supply the whole market. But if tin-plate is made by a large number of firms it will be necessary for them to combine before the way is clear for a single large mill. Combination then brings economies as well as monopoly, and can no longer be singled out as *the* anti-social method of expansion, par excellence.

It was from the railway that this lesson was first learnt. A railway system could not expand competitively, by building a line parallel to other lines, without great waste of capital. After the early years, the only rational method of expansion was by combination. But combination spelt monopoly, and no monopoly was more feared. The obvious advantages of a unified railway system had thus to battle against the Victorian horror of monopoly. A truce might have been called through nationalisation—and this was in fact contemplated in the first railway bills—but private enterprise was held in too great veneration. In the end, safeguards against the

abuse of monopoly power were worked out, and combination was allowed to proceed until all the main line railways were amalgamated into the Big Four in 1921.

The lessons learnt from railway expansion have since been extended to many other industries. When the optimum is large, nothing short of a monopoly or near-monopoly will allow it to be realised. Monopoly is the price which we have to pay for efficiency, if size and efficiency go together. From the social point of view, monopoly, redeemed by economies of scale and regulated by the State, becomes almost respectable; and a reflected respectability falls on combination, in which the motive of economy now mingles with the desire for monopoly. We cannot so easily single out combination for attack when its results cease to be wholly bad, and when exactly the same results might be achieved by the spontaneous growth of a large firm.

The more highly we rate the virtues of size, the better will we think of combination. *Either* method of growth will be more likely to win approval. But combination will be doubly attractive. For when firms are already large, it is difficult and costly to oust them from the market by competition, and comparatively easy and cheap to amalgamate. Especially if the market is narrow, shrinking or uncertain, and if the firms in the field are operating large amounts of durable, fixed capital nothing short of combination will allow firms to grow. The alternative is to build and battle—to duplicate the plant of competitors, and try to annihilate them in a long-drawn-out cut-throat struggle. Thus as the optimum keeps growing, and puts a steady pressure on firms to expand in order to keep pace, it is to combination rather than to plant extension that industry turns increasingly.

## Motives.

**(a) The Economies Motive.**—The first motive to expansion—the prospect of reductions in cost because of economies of scale—has already been discussed. These economies, however, were looked at from the point of view, not of *growth* but of *size*. The angle from which they were approached was not the actual processes by which firms expand, but the comparative efficiency of large and small firms. The optimum scale of production was taken to be within the reach of any firm, and growth towards it from a larger or smaller scale was supposed to raise no special difficulties. But growth does raise difficulties. Expansion towards the optimum is often arrested because the attraction of economies of scale is offset by obstacles to growth. Two obstacles—one marketing, one financial—are outstanding.

The obstacle of a limited market has already been touched on. It is difficult for a small firm to grow to its optimum size if it supplies a scattered market, or if it has to make its way against firms which have already a firm grip on the market—firms, for example, which have won the goodwill of dealers and consumers over a long period of years, or which can afford an expensive advertising campaign if

their market is invaded, or which can bring pressure on dealers to boycott new competitors. This market resistance to growth is enormously stronger when a firm has to expand discontinuously, jumping from a size at which it can supply a local market to a size at which it has to fight for a national market. The two sizes may be "optimum" in relation to the market served, while intervening sizes may raise difficult problems of technique or of management. The risk of failure in passing from the "minor" to the "major" optimum is obviously great, and may deter many firms from sacrificing a comfortable profit to a fifty-fifty chance of death or glory.

Secondly, there is a financial obstacle. A severe setback in trade, against which every firm must guard, puts a very heavy strain on young and small firms, since they have limited access to credit facilities, and no large reserves on which to fall back if they are in temporary difficulties. Firms hesitate, therefore, to push on with schemes for expansion until they have firmly established their financial strength. This hesitation holds up the growth of the firm even when substantial economies are in prospect.

**(b) The Monopoly Motive.**—The second motive—the prospect of monopoly gains—has been identified up till now with profiteering—the charging of excessive prices. We have had in mind the pure monopolist, undiluted by any kind of altruism or by fear of retaliation by other firms. Such a monopolist, with exclusive power over the supply of a commodity, is able to make increased profits, not by rendering some special service or by producing more efficiently, but by selling less and charging more. He can force consumers to put money in his pockets. not because he gives them more, but because he gives them less. He simply exploits scarcity.

But monopoly power may be sought from other motives than the desire to profiteer. First of all, the removal of the threat of competition may, on occasion, make possible genuine economies. When there are two or three firms in a market, each is "kept guessing" by the manœuvres of its competitors, and may spend more time and money in trying to defeat these manœuvres than in improving the services which it offers to the public. Each firm tries to keep in step with the others in public (e.g., by charging the same price), and to steal a march on them in private (e.g., by offering "backhanders," or discounts, or special terms of credit, or supplementary services). Sometimes the public gains from such competition in low prices and better quality, but occasionally it finds itself denied what it wants—a reduction in prices—and offered services which it could well do without. Thus it is offered reams of advertisements in place of cheaper cigarettes, and dozens of milkboys instead of cheaper milk. At the same time, the competing firms are often forced by ignorance of their rivals' intentions into short-sighted policies from which the whole industry must ultimately suffer. They may each instal more plant when prices are high, and then have to cope with the losses in which they are involved by the over-capacity of the industry. A single

firm, with a secure hold on the market, is in a better position to take the long view; freed from the uncertainties of the competitive struggle, it can concentrate on making better and cheaper goods.

Against these economies of unity and stability must be set the temptation to slackness and apathy. A monopoly is like a Government department run by racketeers. It is in a position to do a great deal of good, but it lacks the incentive. There is no need to be enterprising and efficient when a comfortable profit can be made from a cautious dependence on routine. The stronger the monopoly, the less does it need to set its standard of performance against the standard of potential rivals. A firm whose control over any market is constantly challenged by other firms of equal strength must base its control on low costs. It may find competition a nuisance, but it is forced by the "nuisance" to cater for its customers' wants with reasonable efficiency. Remove the nuisance and the spur of competition goes too. A very healthy baby, whose loss would be much lamented, may be lost sight of when the bath water of competition is poured away.

Secondly, firms may unite to form a monopoly in self-defence. They may be afraid of the entry into the market of a new competitor and hope to keep him out by presenting a united front. Two newspapers, for example, may meet a threat to start a third newspaper by amalgamating; in so doing, they take the risk of a fall in profits in order to escape the risk of an even greater fall if the threat to start a third newspaper is carried out. The defensive motive is specially prominent when the market is invaded by a large firm situated in some other part of the country or abroad. Two of the largest British companies—the Imperial Tobacco Company and Imperial Chemical Industries Ltd.—were originally defensive combinations which aimed at meeting American and German trusts on an equal footing.

Self-defence may prompt other kinds of expansion. A firm which has to deal with a monopoly either in buying its materials or in disposing of its products may extend the scope of its operations to strengthen its hand against the monopoly. It may begin to work up its by-products or to manufacture components so as to reap a profit made, partly at its expense, by the monopoly. Sometimes the motive to this extension of scope is not so much self-defence against other monopolies as bargaining strength against competitors. For example, a firm which sells to shops and merchant houses which handle the products of other firms may extend its marketing organisation so as to push its own sales. It may, like many firms in the clothing, boot, chocolate and druggist trades, open its own shops. Or, as in the electrical and engineering trades, it may have its offices or agencies in all the chief cities, abroad as well as at home. In the marketing of tea, petrol, etc., firms very often distribute direct to the retailer, partly from a desire to attach him to their interests.[1]

[1] The same purpose may be served, *without* expansion of the sales organisation, by the grant of an exclusive agency to a single retailer in each district.

Bargaining strength shades into the power to bully other firms. A large enterprise or combine with monopoly powers enjoys comparative immunity from retaliation and can, therefore, take unfair advantage of its business neighbours. It can cancel its orders with far less hesitation than a smaller firm, especially if it is not only a large buyer but the sole, or almost the sole, buyer in the market. It can allow delays in delivery or make delivery of inferior goods. It can threaten to cut off supplies from its customers and force them to give up buying from a potential rival. It may even squeeze out these rivals altogether by organising a boycott, or by bringing pressure to bear on the firms which supply them with raw materials, or on the banks which supply them with credit.[1] Thus the bullying of other firms develops into the exclusion of other firms from the market. Although not protected from competition by its superior efficiency, or by laws granting powers of monopoly, a firm obtains a hold on the market which it is almost impossible to challenge. In theory, any firm is free to compete; in practice, no firm will take the risk. Competition will be particularly difficult if the optimum is large or if the market is narrow. A new firm will hesitate to set up in opposition to established concerns if it can get a footing only by sinking an enormous amount of capital in plant, or in building up a connection. If there is only a limited market, it will be necessary to capture a share from a firm or firms which are already strongly entrenched. The new firm will have to face the task of attracting customers who are tied by inertia, ignorance, or fear to the monopoly. It has to offer not simply a better or cheaper article, but an article so much better or cheaper, or one so effectively advertised that the interest of the monopolist's customers is aroused. It has to make it worth while for these customers to risk the loss of patronage or supplies from the monopolist. It has to hang on through any price-war or boycott which the monopolist may initiate in an effort to smash competition. And all the time the monopolist may be trying to absorb his competitor, or to make terms which divide the market and restore conditions of monopoly in each part of it. Even when a new firm succeeds in invading a market which is monopolised by a large firm, the net change from the point of view of the consumer may in the long run

[1] Competitors may also be exterminated by a price-war, through exclusive dealing agreements, tying contracts, deferred rebates, and so on. Wall-paper manufacturers, for example bind merchants to stock no wall-paper made by outside firms. The Electric Lamp Manufacturers' Association enforces exclusive dealing under threat of boycott, and has virtually ruled out competition through the normal channels of distribution. Shipping rings offer rebates (payable at the end of a stated period) to merchants who undertake not to ship goods by independent lines. Many other trade associations use similar devices. The new firm may also be under the handicap that it has no business connection—no old customers who continue to offer it their patronage in the face of keen competition, no bankers and dealers who continue to offer it credit in bad times.

be precisely nil.[1] The form of monopoly disappears, but the substance remains.

The exclusion of competitors from the market is the very essence of monopoly. Whatever form the monopoly motive assumes, the desire to impede entry into the market is always present. If there is unified control there *can* be no other firms in the market; self-defence is simply defence of a market against invasion by another firm; fighting strength is nothing but power to keep out other firms; even bargaining strength, in the last resort, is based on the threat to exclude some firm from a privileged market or to enter a market previously monopolised. It is this exclusion of competitors which makes profiteering possible. But the danger of profiteering is not always equally formidable. The economies of unified control may even bring prices down; increased bargaining strength over other firms may simply produce a fresh division of the spoils, not higher prices to the consumer. It is only when firms combine for the sole purpose of raising prices, without making any attempt to change or co-ordinate their methods of production, that profiteering pure and simple takes place.[2] If, for example, each firm is approximately of optimum size, they may continue to be separately managed and agree to raise prices in concert. But such outright profiteering is generally diluted by economies both of scale and of unified control when the firms are of less than optimum size, and by the desire to build up a strong combine by stabilising rather than raising prices.

Thus monopoly power and competitive strength are the same thing looked at from different points of view. Both are the product of expansion when expansion loosens the foothold of other firms in the market and makes the entry of new firms into the market more difficult. Strength, whether for bargaining, fighting or throttling, can be used, and is used, to put up prices against the consumer and to push down prices against the supplier of raw materials. It permits increased profits to be made without any improvement in efficiency. The desire for strength or monopoly, while not identical with the desire to profiteer, very often goes with it.

The desire for competitive strength has another side to it—it represents, in Professor Macgregor's words,[3] a "revulsion against risk." There is the risk, as we shall see later, of wide fluctuations in prices, outputs and profits when firms with heavy overhead costs (e.g., railways,

[1] Indeed, prices may *rise*. The volume of business will be divided, so that the large firm will have to produce on a smaller scale, probably at higher cost. If neither firm sees any possibility of driving out the other, a policy of live and let live may be adopted with prices adjusted to the higher level of costs. (cf. E. H. Chamberlin: "Theory of Monopolistic Competition," pages 105-6).

[2] It will be shown later that a sudden increase in demand (e.g., in war time) allows independent and competing firms to profiteer in exactly the same way.

[3] D. H. Macgregor: "The Rationalisation of Industry," *Economic Journal*, 1927.

steamships, etc.) compete in a limited market. There is the risk of failure of supplies or of markets when either comes under the control of a monopoly. There is the risk, too, of recurrent disorganisation of the market by in-and-out firms; these may be of the type which periodically dump a temporary surplus and then retire within their protected market; or they may be floated without adequate knowledge and backing, and meet disaster after a short-lived and ill-advised competition with established concerns. To protect themselves against these and similar risks, firms may seek strength through combination or expansion. No question of profiteering need arise; a reduction in any of these risks is, *so far as it goes*, a gain to society. But the question of profiteering does in fact constantly arise; it is hard to reduce the risk of excessive competition (the first risk cited above), or of intermittent competition (the third risk), without simultaneously raising barriers against competition of any kind. Society gets rid of undesirable competition, but only by sacrificing the main safeguard against the abuse of monopoly power. Once again, strength acquired for one purpose may be used for another; escape from risk may lead straight to profiteering.

The monopoly motive to combination is held in check by various forces of which the chief are public opinion and potential competition. Public opinion makes itself felt through the boycotting of the monopolist's products, or through anti-trust legislation designed to prevent (or at least to obstruct) monopolistic combination. Potential competition is an even more powerful check. If prices are put too high, consumers may turn from the products of the monopolist to substitutes of various kinds. New firms may begin production; old firms in foreign countries or in other lines of business may enter the market; satisfactory substitutes may be devised or may be already on sale. The wider the range of substitutes, actual or potential, the less is the power of the monopolist to charge extortionate prices and the less powerful, therefore, is the incentive to form a combine in the hope of making monopoly profits. If, for example, it is easy for new firms to get a footing in a market, a monopoly will always have the threat of competition hanging over it, and will only be able to maintain its position by superior efficiency. On the other hand, if new firms cannot possibly compete (e.g., because the monopoly holds important patents), it is comparatively safe to charge high prices and draw an inflated profit in "comfortable somnolence."[1] In most trades new firms find themselves in an intermediate position—faced with great, but not overwhelming, difficulties in trying to establish themselves. These difficulties,[2] which are constantly increasing, provide monopolies with a margin when they fix their prices. They can make their prices a little, but not too, excessive.[3]

[1] Report of the Committee on Trusts, page 24.

[2] For examples, see above, page 112.

[3] In comparison, that is, with the price which would satisfy a new firm if it could take over the monopolist's business.

Until comparatively recently the threat of new competition was particularly powerful in Great Britain. The natural resources of the country were plentifully scattered and difficult to monopolise, so that new firms could not readily be deprived of supplies. Foreign producers were within easy reach of the British market, and there were no tariff barriers to keep them out; British producers, therefore, had to face the competition of imports, and so long as this was so, they could not easily exploit the home market. It was impossible also for one firm to obtain exclusive use of the means of transport, or to secure special treatment from the railways (like some of the American trusts); goods could be shipped cheaply by land or by water, and as distances were very short the cost of transport was rarely a large or strategically important item. With the adoption of a high tariff policy one of these safeguards against monopoly has gone. At the same time the growing importance of large-scale production strengthens the established firm. To found a small firm is easy, but to found a large firm in a market which is no longer expanding rapidly is a very risky undertaking. The hold of the established firm over the channels of distribution is also becoming steadily tighter, and the task of the new and growing firm correspondingly stiffer. Finally, the State, instead of obstructing, is now compelling, combination and has brought into existence a host of monopolistic organisations in agriculture and fishing, coal, steel, cotton and other industries. This fostering of monopoly in some industries will inevitably have its repercussions in others. The monopoly motive may play a bigger part in British industry in the future than it has in the past.

(c) **The Power Motive.**—The desire for size and strength may have its origin in a striving after power rather than after profit. Thus, in addition to the two motives to expansion which have so far been discussed, a third—the desire for economic power for its own sake—very often enters. A business is an instrument both for the making of profits and for the winning of personal power and a place in the sun. Its growth is governed, therefore, almost as much by the pride and ambition of its owners as by the prospect of savings in costs or monopoly gains. Control of a large business is flattering to a man's sense of importance; he has scope for his energies, and freedom to plan and speculate and build after his own heart. His work is stimulating, because responsible and creative. It carries with it power and leadership over large numbers of workers; the excitement of a game; the satisfaction of a rôle in life; a sense of real achievement; the prospect of founding an industrial dynasty. These are incentives to which even the salaried administrators of joint-stock companies respond. Their eagerness to extend the operations of their company is shown both by the large sums which they put to reserve instead of distributing in dividends, and by the frequency with which they overreach themselves in trying to absorb other companies. The motive to expansion either out of undistributed

profits or through amalgamation is very often power rather than prudence or profit.

Economic power carries with it independence. But love of independence by itself is an obstacle to growth. In a limited market, if one business is to expand, others must be pushed towards bankruptcy or must combine with the first. But if independence is highly valued, each firm may prefer to hang on stubbornly in spite of low profits rather than merge with an expanding competitor. Those who put independence before profit probably outweigh those who are willing to sacrifice profit to ambition, and on balance, therefore, the desire for economic power prevents rather than promotes industrial combination and the growth of the firm.

**(d) The Financial Motive.**—When a firm expands by combining with other firms there is a profit to be made by *effecting* the combination quite distinct from the profits of the combination *once effected.* Hope of making a profit of the first kind is still another motive of expansion. How does such a profit arise?

Industrial combination is generally financed through the raising of capital on the Stock Exchange. The raising of capital is the work of professional "financiers" (usually bankers), and is undertaken by them either on their own initiative (if they are "promoting" the amalgamation), or at the request of companies which have agreed to amalgamate. The financiers who "promote" industrial combinations (i.e., arrange them "off their own bat," from purely financial motives) take a special interest in companies whose prospects can be made to seem very attractive—more attractive than well-informed and cool-headed persons (the financiers themselves, perhaps) judge them to be. Shares in the companies can be sold at handsome prices to the investing public and so net correspondingly handsome profits for the financier making the issue. Financiers, in short, trade on the optimism of the investing public and are always on the lookout for bait. There is no better bait than an amalgamation. No one, not even the financiers, can say with much hope of accuracy what the real prospects of the combination are. But it is possible to make great play with the various economies of large-scale production from which the combination is expected to benefit, the reputation and business connections which have been built up by the undertakings which are being merged, the extensive and growing market which they serve, and so on. The earning power of the combination can then be capitalised very generously so as to set a value on the assets far in excess of their cost, and investors, if in the right mood of speculative frenzy, will swallow the bait and buy the shares.

The opportunity of making large profits from industrial consolidations is a direct incentive to financiers to promote such consolidations even when they have little real justification. Combines are organised which have only a faint chance of success, but which bring enormous profits to the original organisers. So long as all the shares have not been disposed of, the financiers use their control over the combine

to preserve an appearance of profitability and distribute dividends which have not been earned. Once their holdings have been disposed of they are indifferent to the fate of the company. Very often failure appears to be due to lack of working capital and inability to obtain adequate credit from the banks. But in reality the crisis has its origin in the use of working capital and bank credit in the past to make interest payments to bondholders.[1]

When the combine's prospects of ultimate success are good, it may survive this financial piracy. A well-known example is the United States Steel Corporation, now a flourishing concern producing from 30 per cent. to 40 per cent. of the output of the entire American steel industry. When the corporation was formed in 1902 the profits of the promoting syndicate were no less than $62.5 millions, and of this sum J. P. Morgan & Co. alone received $12.5 millions as syndicate managers, together with a further share as syndicate members.[2] At first the company had difficulty in paying dividends, the average dividend on the common stock (or ordinary shares) being less than 2 per cent. over the first nine years.[3] Later, however, the reinvestment of profits on an immense scale improved the earning capacity of the company and it was able to pay high dividends until the coming of the depression in 1929.

Other combines have been less fortunate. Professor Dewing, after studying the history of fourteen large combines which underwent reorganisation in pre-war America, came to the conclusion that "had the early interests been willing to forego immediate profits and conserve the funds of the corporations, the enterprises could have been placed on a sound footing and the men themselves have secured vastly greater returns."[4] The financier tends to go for a quick profit and a large one, rather than nurse a company over a long period or finance a series of small but fairly profitable developments. Where the financial motive to expansion is the dominant one, therefore, we tend to find a mushroom growth followed by collapse. Where the financial motive is absent, as in the slow and piecemeal extension of a new line of business, or in the unspectacular combination of small firms, growth is more steady and secure.

(e) **Other Motives.**—To these four motives to expansion might be added others of less importance. For example, a company may be forced to expand in order to comply with changes in the law. It may require either to add a new department or extend an old one or combine with some other firm. An Act requiring the compulsory pasteurisation of milk would be likely to affect dairymen in one or other of these ways. Again, there is sometimes what might be

[1] A. S. Dewing: "Corporate Promotions and Re-organisations," page 555.

[2] H. A. Marquand: "Dynamics of Industrial Combination," page 76.

[3] *ibid.*, page 112.

[4] Dewing, *op. cit.*, page 550.

called a "fiscal" motive, when the exemption of a company's undistributed profits from tax puts a premium on expansion.[1] The shareholder has the alternative of seeing £1 put to reserve or receiving, say, 15s. in dividend and investing it in some other company. Naturally he will be inclined to plump for the first alternative and approve—if his approval is sought—the building up of reserves for use in making extensions of plant or in buying up other companies. On the other hand, if profits put to reserve are subject to tax, an obstacle to expansion is created. The young and efficient firm which depends on high profits to provide capital for growth takes longer to overhaul the established firm which has already more capital than it can use.[2]

**Method: Growth and Concentration of Control.**

A firm may grow either by extending its plant or by combining with other firms. The first method involves an expansion in the capacity not only of the firm, but also of the whole industry; the second method produces a change in the pattern of ownership and control of the industry, but not in its capacity. Of the two, the second is much the more complex; combination covers a wide variety of methods of growth.

When we speak of two firms combining we generally think of two quite independent units coming under the same management. But how many firms nowadays are "quite independent"? The management of one firm nearly always has a finger in the pie of other firms. Perhaps the directors have seats on the boards of other companies; or influential shareholders are equally influential shareholders in other companies; or friends and relatives of the directors, the shareholders or the staff hold positions of responsibility in other companies. Every company has personal links with others and these links limit its independence. The management is forced to show regard for interests outside the firm, and that regard may undergo little change once the forms of combination have been gone through. The same staff, the same shareholders, the same directors—even the same *policy*—may remain after "combination." What is generally called "combination," in short, may be no more than a tightening of the links between two firms with common interests—more a gesture of recognition of mutual dependence than a destruction of independence.

[1] Such exemption does not exist in Great Britain, but the proposal for exemption has often been made.

[2] In some industries, however, a premium on expansion may persist even when all profits are subject to tax. Various outlays which help the business to expand can be charged, by judicious accountancy, to "income account" (i.e., treated like other items of ordinary expenditure), and so escape tax, although the company benefits from such outlays in exactly the same way as from profits put to reserve. For example, no tax is paid on advertising outlay, but advertising helps to build up an important asset in the form of "goodwill."

A firm's independence or freedom of action may be limited not only by personal links, but also by trade etiquette or by understandings with other firms. In the eighteenth century, Adam Smith observed that "people of the same trade seldom meet together even for merriment and diversion but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices."[1] Nowadays "conspiracies against the public" are as common as ever. In a large number of markets the prices to be charged and the areas to be supplied by each competitor are settled by gentlemen's agreements and informal understandings. In these understandings trade associations play an important part. By bringing business men into close personal contact, or by promoting the exchange of information on methods of cost accounting in use, prices charged, output, costs and so on, trade associations create an atmosphere of co-operation favourable, first to policies of "live and let live," and later to joint action to maintain prices and restrict output. The information collected and exchanged provides a basis for a common policy and suggests the need for one. Very often, however, such a policy is narrowly conceived in terms of monopoly at its worst. It is designed to put an end to price-cutting and protect profits on invested capital. It does this, but only at the price of perpetuating inefficiency. There is neither the comprehensive planning and general overhauling of technical methods that might be expected to result from outright combination, nor the weeding out of inefficient firms and inefficient methods that might be expected to result from competition *à outrance*.[2] Instead, pressure is put equally on efficient and inefficient firms; the efficient firms are clamped down within their existing markets while the inefficient firms survive freely in theirs.

Understandings do not amount to combination in the narrow sense. But they represent a sacrifice of independence and a limitation of competition which are comparable in their effects with actual combination. Combination in the sense of concentration of control is a matter of degree rather than of kind, and understandings between competitors carry us a stage nearer to full concentration.

Sometimes firms come to more formal agreements, fixing minimum prices. Practically every British industry of importance has had experience of such agreements at one time or another. Examples are the fixing of minimum prices for milk by the Permanent Joint Milk Committee and later by the Marketing Boards, for cotton yarn by the Federation of Master Cotton Spinners' Association, for tramp shipping freights by the Tramp Shipping Administrative Committee,

[1] "Wealth of Nations" (Ed. Cannan), I, page 130. A modern illustration is provided by the famous dinners (1907-11) to which Judge Gary of the United States Steel Corporation used to invite rival manufacturers so as to exhort them against price-cutting.

[2] As will appear later, competition does not always make for the survival of the efficient firm. The old and well-established (but inefficient) firm may ride the storm of trade depression more successfully than the young and efficient (but poorly connected) firm.

for wireless valves, electric lamps, steel rails, cast-iron pipes, and many other electrical and metal products by various trade associations. The fixing of minimum prices is designed to prevent price-cutting. But whenever the temptation to price-cutting is great (e.g., under the stress of trade depression), the agreement tends to be evaded or openly violated. It can be evaded by the offer of high discounts or of special concessions such as a low price for articles which do not come under the agreement. And it can be violated with impunity because it is a contract in restraint of trade and is, therefore, not enforceable at law.[1]

Formal agreements to share the market are also common. Shipping conferences, for example (i.e., associations of steamship lines), arrange for the division of traffic by limiting each line to a certain number of sailings or to certain ports of sailing. The market in thread has been shared by agreement between two large combines—J. & P. Coats and the English Sewing Cotton Co.—for over thirty years.[2] In international trade there are agreements to share the market in arms (between Vickers-Armstrong and Du Ponts); in chemical products (between Imperial Chemical Industries, I.G. Farbenindustrie, and the corresponding companies in France and Switzerland); in tobacco products (between the Imperial Tobacco Company and the American Consolidated Tobacco Company); and in electric lamps, gas mantles, and numerous other products. When competition is by tender, arrangements may be made to share the market through an association which decides who is to make the lowest tender and take the contract. The other members either do not tender at all or put in higher tenders as a blind. Sometimes sharing the market is combined with a pooling arrangement. To take shipping again, part of the freight on some cargo—say, tea from Calcutta—may be pooled and divided in an agreed ratio between the conference lines engaged in the trade. Some of the lines trading from Calcutta may carry little or no tea, but will nevertheless participate in the pool. An alternative arrangement is for each line to be allotted an agreed proportion of the traffic and to pay a fine into the pool on cargoes in excess of this proportion, or draw from the pool to compensate for any deficiency.

Agreements to fix prices or to share the market are very often combined with agreements to limit output. Indeed, prices cannot be kept up in the face of falling demand *unless* producers cut down their output and offer a smaller quantity of goods for sale; if producers continue to press supplies on the market, prices *must* fall below the "fixed" minimum, or the supplies will remain unsold. Similarly, sharing the market means firms bind themselves to limit output to

[1] In the United States "restraint of trade" is interpreted more strictly than in Britain, and in Germany much less strictly.

[2] In Germany, the brewers not only undertake not to entice away each other's customers, but actually pay a fine if any clients come to them voluntarily from some other brewery. (R. Liefmann: "Cartels, Concerns, and Trusts" page 32).

the needs of their own (reserved) market or to some fixed proportion of total trade. Restriction of output is implied, therefore, in agreements to share the market or to refrain from price-cutting; the responsibility of limiting its output falls on each firm individually. Sometimes, however, associations are formed to organise restriction of output for the industry as a whole. When an industry is depressed, for example, each firm may agree to work only a proportion of its plant. Or the aggregate output to be produced may be fixed by a central body and divided up between the members of the association in proportion to their output in the past. Each firm is allotted a quota. If this quota is exceeded, a fine must be paid into a pool in proportion to the excess; while, if a firm fails to reach its quota, a payment is made to it out of the pool. When this pooling system is not in operation, it is generally possible for producers to expand output beyond the limit set by their quota by purchasing the unworked quota of other producers. Later, when quotas are under revision, these purchases of quota can be used as a lever to procure a higher share of output under the new agreement.

Associations which not only fix prices or allot quotas, but also undertake the business of marketing, are called cartels.[1] A cartel acts as a selling agency on behalf of its members, distributing orders between them in accordance with the same agreed formula, and not interfering in any way with their internal management. The functions of producing and of marketing are separated, production being carried on by independent units under semi-competitive conditions, while all sales are made through the agency of a monopoly. The profits, or losses, of the sales monopoly are pooled and shared between the members in proportion to output. Sales in special markets (e.g., abroad) may be made at prices varying with the keenness of the competition that has to be met. Prices in the home market, for example, may be kept up while sales at cut-throat prices are made with a view to driving out competitors in some foreign market. Or a special levy per unit of output may be made on each firm so as to subsidise exports.

Until recently the cartel form of organisation hardly existed in Britain, although a cartel of the modern type—the Newcastle Vend—was formed in the British coal industry as far back as the seventeenth century.[2] Cartels were much commoner in Germany, where the law

[1] The term is frequently applied to some of the looser types of association described above. In the strict sense, however, a cartel means primarily a selling agency with monopoly powers acting on behalf of independent producers.

[2] H. Levy: "Monopolies, Cartels and Trusts," page 106 *et seq.* This cartel, the first of its kind, had an elaborate statute binding its members in a very strict way. It allotted quotas, which were revised from time to time, and entered into agreements with the London wholesale coal organisations. The cartel did not break down until 1844 when the development of railway transport put an end to its monopoly of the London market.

put no obstacles in their way. There were cartels for pins and for push-buttons, for soap and for salt, for artificial palm trees, merry-go-rounds and roller-skates, for toilet paper and even for special kinds of toilet paper.[1] In Britain a number of central selling agencies, resembling cartels, have been set up since the last war—most of them by the State itself. The Milk Marketing Boards, for example, have practically a monopoly of the milk supply, and act as the agents of the dairy farmers in arranging terms of sale to distributors of liquid milk and to manufacturers of dairy products. The Boards fix different prices for milk according to the uses to which it is to be put, and pool the proceeds, distributing them amongst the farmers who supply the milk.[2] Thus the profit on milk sold in the liquid market is set against the burden of loss on milk sold in the manufacturing market (for manufacture into butter, cheese, etc.), and is shared over the whole mass of producers. At the same time the producers are left free to manage their farms as they please without even the restriction of a quota.

The cartel, and the looser types of association which have been described so far, all stop short of actual combination. They are all "terminable associations" as opposed to "trusts."[3] A trust is simply a very large firm formed by consolidation of independent companies, or a group of associated companies under the control of a single interest, the firm or group of firms being large and strong enough to exercise powers of monopoly. Trustification may take place in a variety of ways by amalgamation, by acquisition of a controlling interest, by the formation of a holding company, and so on. The biggest of the British trusts—Imperial Chemical Industries—was formed by the issue of shares in exchange for those of four large companies—Brunner Mond's, United Alkali, Noble Industries and the British Dyestuffs Corporation—each of which was in a strong position in one branch or other of the chemical industry. Another example of a British trust is Lever Bros., which grew by absorption of competitors until it was producing 75 per cent.

[1] R. Liefmann, *op. cit.*, page 30. There were perhaps as many as 3,000 cartels in German industry and trade in 1925. In the last few years so many new cartels have been set up that practically the whole of industry is now cartellised.

[2] The pool is augmented by levies on farmers who do not sell through the Board but make retail sales of milk on their own (i.e., sell exclusively in the high-price market). On the other hand, a subsidy is offered from the pool to farmhouse cheesemakers (i.e., on milk which fetches less than the pool price). Various other items (administrative expenses, premiums on special grades of milk, etc.), are also deducted before calculating the pool price.

[3] An intermediate stage between terminable associations and trusts is reached when one company agrees to pool its profits with another or to lease to it part of its plant.

of the soap made in this country.[1] The subsidiary companies in a trust of this size generally retain some real independence, and only the broad lines of policy - especially price policy—are dictated to them. If the subsidiaries are scattered up and down the country, centralisation will be particularly difficult. Where local goodwill is important (e.g., in banking) the outward appearance of independence must be preserved; and, where local conditions vary, the substance of independence must often be preserved, too, for the central office will hesitate to take the responsibility of interfering vigorously in the management without adequate knowledge.

When these obstacles to centralisation exist, trusts and cartels may differ little in their results, however much they differ in form. Each of them becomes simply a means of combining bigness in marketing (for the sake of economies of large-scale marketing or, more probably, for the sake of monopoly and competitive strength) with smallness in producing (for the sake of efficient management).[2] When the obstacles to centralisation are removed—i.e., when the managerial optimum expands—a real difference emerges between the trust and the cartel. The cartel is then a prop for units which are uneconomically small; trustification, on the other hand, is a means of bringing units together into a firm of something like optimum size. Like the cartel, the trust can integrate the price policy of the units, enforcing uniform—and generally higher—prices; unlike the cartel, it can also integrate their methods of production, suppressing inefficiency, specialising plants to particular products or markets, standardising varieties of product, and using wide powers of co-ordination over the whole field of production, research, marketing, and so on.

Combination, then, ranges from loose agreements at one end to outright amalgamation at the other. At each stage control over price and over methods of production becomes more and more concentrated. First the monopoly powers of the combine are consolidated through more stringent agreements between the members; then the unit of management is expanded through the trustification of the members. Which method of combination is selected depends upon the advantages and costs of bringing about concentration of control.

Let us take the advantages first.[3] Full concentration (e.g., in the trust) is to be preferred whenever economies of scale are in prospect. The optimum unit of management, for example, may have grown larger, or technical changes may be making for expansion in the size of the business unit. Mere understandings and agreements generally do little to adapt industry to these changes, but tend rather to hold them up. Secondly, where the interests of producers differ widely, it may be hard to arrive at agreement between them short of amalgamation. Even if a looser agreement is possible, it may have

[1] L. Fitzgerald: "Industrial Combination in England," page 62.

[2] Or because producers and managers have an intense desire for independence.

[3] Advantages to the combine, not necessarily to the public.

to be too vague or too exposed to the risk of disloyalty to be of value.[1] Consolidation in a trust, however, will ensure unity of control and a common policy. The vested interests which block the concentration of production in the most efficient plants will find the ground cut from their feet, and at the same time each plant can be forced to charge a uniform scale of prices. Thirdly, full concentration may be the most effective method of suppressing competition. So long as firms are linked loosely together by agreements or understandings, they may find it harder to take concerted action against a new competitor than a large trust. Lever Bros., for example, are in a far stronger position in the soap trade than any cartel of soap manufacturers that might have taken their place. Fourthly, consolidation can be extremely profitable to those who promote it, whereas it is impossible to "cash in" on a trade agreement or even a cartel. In addition, it is generally easier for a trust than for a cartel to raise new capital when an industry needs to be re-equipped and re-organised. The new capital may be raised from the public when the trust is formed, or may be supplied by firms with surplus reserves when they are merged with firms in financial difficulties. Finally, the law generally puts more obstacles in the way of the cartel and agreements "in restraint of trade" than in the way of outright combination to form a trust.

Most of these advantages must be qualified, however. First of all, the economies of scale which can reasonably be expected from consolidation are often grossly exaggerated. The units consolidated, as we have seen, may continue to be managed much as before. Or the trust may easily pass beyond optimum size and be too unwieldy for efficient management. This is in fact one of the most common reasons for the breakdown of trusts.[2] Secondly, it is not to be supposed that disunity dissolves into unity whenever the wand of consolidation is waved. Sectional interests persist inside the trust, pushing the claims of one plant against another just as they

[1] Compare, for example, the following comment on the working of the Coal Mines Act, Part I: "Collieries with newly opened seams have bitterly resisted the regulations which have prevented them from expanding. The exporting colliery has demanded prices that would enable it to meet foreign competition, prices which would be unprofitably low for the inland colliery. One district has insisted upon uniform prices at the point of consumption; another has declared for uniform prices at the pit-head. The mechanised mine with heavy fixed costs has desired low minimum prices and continuous operations; others have preferred high prices and part-time operations. These are but a few of the many questions that have continuously threatened to disrupt the industry." (A. F. Lucas: "Industrial Reconstruction and the Control of Competition," page 100.)

[2] cf. Dewing: "Corporate Promotions and Re-organisations," page 558. It is too often assumed that the profits of a trust cannot be less than the joint profits of the firms out of which it was formed. It has repeatedly happened that earnings after consolidation have worked out far below previous earnings, in spite of the trust's monopoly advantages and in spite of the optimistic forecasts of promoters. (Dewing, *op. cit.*, page 547).

might in a cartel, but generally with less vehemence and certainly with less power to threaten mischief. Thirdly, the strength of a trust is not necessarily greater than that of a cartel. A trust rarely controls the whole industry, whereas cartels very often do. The looser form of combination is the more elastic and comprehensive, and elasticity and comprehensiveness are often more essential to strength than unity. The monopoly of the salt trust—the Salt Union—was twice broken by new competition, but re-established successfully when it formed a cartel with its chief competitors. Finally, legal obstacles to trustification are sometimes at least as overwhelming as legal obstacles to other methods of combination. For the Big Five in banking or the Big Four in railways to combine without Parliamentary approval and encouragement would be impossible. But they are free to make agreements fixing minimum rates of interest on loans or to enter into pooling arrangements for traffic.

These qualifications are reinforced by the comparative costliness of forming a trust. There are, first, the costs of arranging the merger itself. These include not only lawyers' fees and bankers' commissions, but also the financial profits of the promoters. They include, too, excessive payments made to manufacturers for their plant—the cost, for example, of old and inefficient plants which are bought up only to be closed down almost at once; and the cost of firms which earn little or no profit for the trust, but which hold out successfully for exorbitant compensation because they have a high "nuisance value"—power, if left outside, to make things awkward for the trust. Secondly, there are the costs of meeting new competition attracted by the formation of the trust.[1] Manufacturers who have sold their out-of-date plants to the trust may invest their money in new factories, modern in design and the last word in efficiency. They can draw on long experience in choosing a site and machinery for their plant, and on thorough familiarity with trade conditions and the special wants of influential customers in managing their business.[2] Thus they are formidable competitors of the trust—so much so that an undertaking not to compete for a period of years is often asked for from producers when they are bought out. Other, less experienced, competitors may be attracted by the high profits which the promoters of the trust anticipate.[3] And there will also be strong competition from firms not absorbed

[1] Such competition will be impossible if the trust has a monopoly of some essential raw material, a stranglehold on the channels of distribution, or legal protection through legal patents or franchises. In all other circumstances, competition will be possible, but not necessarily easy (see above, p. 112).

[2] cf. Dewing, *op. cit.*, page 564.

[3] It may even happen that business men have factories put up for the express purpose of inducing trusts to buy them out. In the early history of American railways, for example, branch lines were sometimes built and offered for sale to the main line company on exorbitant terms.

by the trust—all the stronger because the burden of regulating output and stabilising prices will fall heavily upon the trust.[1] To get rid of all this competition, trusts are often forced to expand, buying up one after another of their competitors at great cost, but without putting an end to new competition. Alternatively, they have to abandon hope of exploiting their monopoly and concentrate their energies on securing the economies which size brings within their reach.[2]

In comparison with the cost of forming a trust, a cartel is inexpensive to organise and the cost of framing an agreement between trade competitors is negligible. A cartel, it is true, may be embarrassed by new competition much as a trust is, and may be forced to make terms (e.g., the grant of a large quota to new competitors) which are just as costly as purchase at an exorbitant price. But this new competition is not brought into play more or less automatically through the formation of the cartel; there is no displacement of managerial talent and capital, less publicity, and less need to shoulder the burdens of other producers by allowing them to share in the advantages of a stable market. The new competition which a cartel has to face is the penalty either of inefficiency or of exorbitance.

The view which producers (or promoters) take of these advantages and costs determines how far combination will go and what form it will assume. If the advantages are great (e.g., if important economies are likely to result) and if the costs are low (e.g., if producers are willing to surrender their independence cheaply) amalgamation will proceed rapidly. If the advantages are limited (e.g., if each firm is already of nearly optimum size) and if the cost of outright combination is high (e.g., if it is difficult to promote a trust because investors will not "bite"), then resort may be had to some other form of combination more in keeping with the special advantages hoped for, and the special obstacles met with. A wide variety of methods of combination may be considered, some concentrating control over price, some concentrating control over production also. One method of combination may be found which sets up a pattern of control superior to the existing pattern.[3] If the cost of changing from one to the other is not excessive, it will be adopted. It is this change in the pattern of control that we mean by combination.

[1] The smaller steel companies in the United States, for example, seem to have gained because the United States Steel Corporation carried the burdens of leadership.

[2] In the end, this may put the trust's monopoly far more securely beyond challenge than any patched-up combination.

It should be noted that much of what is said above about the advantages and costs of combining to form a trust does not apply when combination takes place between two firms not aiming at monopoly.

[3] The superiority may arise from a concentration of control over the amount marketed and the price charged, or from a concentration of control over methods of production, or from a blend of concentration of the one with dispersion of the other.

Thus we return to our starting point. Firms are not to be thought of as atoms which now and again coalesce. They are to be thought of in biological terms as organisms that grow and decay. Their growth is conditioned by the play of opportunity on a given endowment of managerial ability, financial strength, and so on, just as the growth of organisms is conditioned by the play of environment on a given hereditary endowment. Firms grow by grafting and proliferation rather than as homogeneous units with which new units of the same kind are geared mechanically.

### Direction: Growth and Integration.

The growth of a firm, as we have seen, may take place by either of two methods—plant extension or combination. Cutting across this division is another, based on the *direction* of growth. When a firm increases the size of its establishment, or combines with neighbouring establishments of the same type, it may continue to make the same products by the same processes on the same site. Generally, however, the growth of a firm, whether it takes place by plant extension or by combination, involves changes in the site or in the scope of the firm's operations. The term "integration" is applied to changes which add new products and processes and the term "disintegration" to changes in the direction of fewer products and processes.[1] The use of a larger number of sites, when the firm builds or buys branch factories in other parts of the country, is called "diffusion"; and the enlargement of one establishment accompanied by the closing down of establishments in other parts of the country is called "concentration."[2]

A firm may grow, therefore, along one of several routes. It may grow *horizontally* by combining with firms which make similar products; *vertically* by undertaking processes of manufacture in continuance of those which it already performs; *laterally* by extending the list of products which it turns out; *territorially* by operating over a wider area. But growth is rarely one-dimensional—it takes place, as a rule, along several of these routes simultaneously. A firm which sets up a branch factory, for example, grows horizontally if the new factory produces goods similar to those manufactured in the main establishment; vertically, if the new factory runs a repair

[1] No attempt is made in this section to distinguish carefully between integration as process of growth and integration as a state of affairs. For example, a firm may add new products and simultaneously give up making others, leaving fewer products than before. The distinction is generally not of importance and when it is, attention is drawn to it.

[2] Changes in the direction of more or fewer sites might be called territorial integration and disintegration respectively, so as to preserve the parallel with changes by product or process (see above, page 57). Unfortunately these labels have come to be applied the other way round, territorial integration, for example, meaning concentration on a *smaller* number of sites. To avoid confusion, therefore, the terms "concentration" and "diffusion" have been used in the text.

department and the main establishment does not; laterally, if the range of products is slightly different; and territorially, if the branch factory is at a distance from the main establishment. Thus the pattern of growth is often woven from different types of integration. Sometimes it is woven, not from integration alone, but from integration mixed with disintegration. A firm may decide, for example, to increase the scale of output of standard lines (horizontal integration) and cease production of special lines (lateral disintegration). Or it may decide to abandon some processes of manufacture (vertical disintegration) so as to specialise on a larger scale on the remainder (horizontal integration). Or again, it may add a new stage of production at one plant (vertical integration) and close down a similar plant elsewhere (concentration plus horizontal disintegration).

(a) **Horizontal Integration.**—Horizontal integration leaves the scope of a firm's activities unchanged. It may take the form of an extension of plant and an accompanying increase in output without change of product or process; or, alternatively, it may consist of the combination of firms making similar products. For example, an iron-smelting company may build more blast furnaces or combine with another iron-smelting company.

The normal method of growth is horizontal. The firm which is successful in one line of business naturally seeks to extend that line. The firm which is anxious to protect itself against price-cutting and loss of business tries to combine with other firms in its own trade. It is not only that business men like to stick to a trade with which they are familiar. Both the economies motive and the monopoly motive press far more powerfully horizontally than in any other direction. If a firm grows by widening the scope of its operations rather than by simple expansion of output, economies of scale must as a rule be sacrificed. Similarly, a combine of miscellaneous firms generally lacks the strength of a combine of similar firms. Thus it is only if there are special obstacles to horizontal integration that growth takes place in other directions.[1]

(b) **Vertical Integration.**—Vertical integration is the union of a sequence of processes formerly carried on by separate firms. Three different varieties of vertical integration can be distinguished. First, an extension of the process of manufacture, backwards towards the raw materials or forwards towards the market. For instance, a steel firm may take over the previous stage of production and build its own blast furnaces, or combine with a company which is engaged in

[1] It is unnecessary to give illustrations of horizontal integration since *any* growing firm provides an example. Large horizontal combines can be found in practically every British industry (e.g., Wall-paper Manufacturers' Association which controls over 90 per cent. of the wall-paper industry; the Distillers' Company, controlling over 80 per cent. of the whisky trade; J. & P. Coats; Allied Ironfounders; Fine Cotton Spinners, etc., etc.).

the production of pig-iron ("backward integration"); or it may continue processes already performed and build rolling mills or amalgamate with the firms which buy its steel ("forward integration"). Second, auxiliary goods and services required in the manufacture of a firm's main products may be provided within the firm instead of purchased from outside. For example, the firm may undertake its own repairs, generate its own power, or make its own tools or designs.[1] Third, the special service of marketing may be taken over by the firm.

The motives to vertical integration differ with the type. The motive to forward integration, for example, is generally to find a market; the motive to backward integration is to secure sources of raw materials. Forward integration is not uncommon in times of depression, when firms are anxious to push the sale of their product and so reap the economies of operating at full capacity. Backward integration is to be expected mainly in times of boom when there is a danger of a shortage of raw materials. Thus a steel firm might buy over a shipbuilding yard when trade is bad, and blast furnaces when trade is good. Since combinations of all kinds are easiest to carry through in times of boom—the financial motive being very strong at such times—the antithesis is never quite so sharp as this. But there is a *relative* preponderance of forward over backward integration in times of depression, and of backward over forward integration in times of boom.[2] Similarly, the assumption of marketing functions is generally prompted by the desire for bargaining strength,[3] whereas the motive to provide auxiliary services within the firm is more often to be found in the technical economies of linked processes.[4]

Of the economies of scale to which vertical integration leads, two only are of much importance. The first—economies of linked processes—have been particularly striking in the steel and chemical

[1] P. S. Florence, *op. cit.*, page 22. Professor Florence gives the name "diagonal integration" to this type of expansion.

[2] Whether combination is forward or backward must be judged from the motives of the dominant partner.

[3] See above, page 73. The technical difficulties involved in integration of production with marketing are very great, unless the retailer handles only a narrow range of goods. The manufacturer of shovels, for example, will have no wish to retail them if he has to handle simultaneously everything from hot-water fittings to hearthrugs and even goats (cf. H. Smith: "Retail Distribution," page 94 n.). Nor will the perplexities of drapers who sell cat's meat and ladies' underwear in Bethnal Green (*ibid.*) appeal to the textile manufacturers of Lancashire. Even when retailing is highly specialised (e.g., in shoes, oil products, motor-cars, etc.), it is rarely possible for the producer to defend his entry into retailing on grounds of efficiency. The real motive is generally the bargaining advantage which he obtains by building up a special market for his goods.

[4] See above, page 66.

industries, where it is essential to conserve heat by having each process in close conjunction with the succeeding one. In other industries, it is economical to link auxiliary services with the main process or processes (e.g., the making of machines, the generation of electricity, etc.). Generally speaking, however, it tends to be more economical to split up processes than to link them, since splitting (disintegration) allows specialist firms to carry on each process on a large scale. It is to the other set of economies—economies due to the spreading or avoidance of risk—that we have to look for the main force working against disintegration and in favour of integration.

The chief risk against which vertical integration safeguards a firm is failure of supplies. A firm which is unable to count on regular and punctual delivery of raw materials runs the risk of being constantly held up by lack of supplies. It can protect itself against this risk by holding large reserves of stocks, but this procedure is both costly (since there are storage and interest charges to be met), and risky (since the price of stocks may fluctuate enormously). The firm may be forced, therefore, to secure direct control over the supply of its raw materials through vertical combination. Again, a firm may make its own materials in order to be certain that they are of good and uniform quality. Or it may do so because of the risk that supplies will be unprocurable before its physical plant has had time to wear out, or before a full return has been obtained on the cost of building up the business organisation and sales connections. Or a firm may undertake the manufacture of its materials as a precaution against a rise in their price, or against the danger that they may come under the control of a combine, and be monopolised. Vertical combination, in fact, is very often a counterstroke to horizontal combination.

An interesting example of vertical integration, which illustrates the influence of some of these risks, is the control of over 70 per cent. of their ore requirements by British pig-iron producers. So long as the average ironworks remained small, and producers had free access both to scattered deposits at home and to imports from abroad, integration made little progress. When the unit of production grew large enough to be worked in conjunction with a coal or iron mine the tendency to combination began to assert itself. The ironworks had a reliable supply of materials of suitable and known quality, and the mines had an assured market with less risk of irregular operation. Further increases in the scale of production made a large and steady supply of ore imperative. This accelerated the process of combination, for British ironmasters, to be sure of a supply which would be regular and lasting, bought up deposits both in Britain and in other countries (e.g., in Spain). The motives to the purchase of foreign iron mines were particularly strong. Foreign deposits were more concentrated and therefore more suited to the needs of large ironworks than the scattered British deposits. The danger of monopoly or of control by a competitor, just because of this

concentration, was correspondingly great. During and after the war, therefore, many of the large British iron and steel firms took steps to assure themselves of a regular supply of ore by buying foreign iron mines.[1]

A vertically integrated firm enjoys various other economies. It can effect economies of unified control of the succeeding stages of production. If each stage is in separate ownership, the chain of independent firms may hang back from innovations which are in the common interest of the group, but from which the innovator derives little advantage. The design, quality and, above all, the durability which one firm looks for in its components may be unattractive from the point of view of the firm supplying them. Integration enables the conflicting interests of the various stages of production to be reconciled; each firm's special knowledge, and the profits which result from its use, can be pooled in the combine and new processes of common advantage can be introduced. Moreover, a policy which keeps each stage of production in step with the one before and the one after can be formulated. In the cotton industry, where there is little vertical integration, a change in demand must be transmitted gradually along the chain of independent firms, each of which, reacting in ignorance of what is happening further along the chain, easily misinterprets market tendencies and makes an excessive adjustment to the change in demand. Each firm becomes progressively more alarmed or more hopeful than the one in front until the repercussions are quite disproportionate to the original change. In motor-car production, on the other hand, each department is synchronised with the others, from the purchase of raw materials to the marketing of the finished product, and adjustment to reports by retail agents of a change in demand can be made quickly all along the line. Integration introduces more comprehensive planning.

Vertical integration is often a reaction to monopoly or the threat of monopoly. We have seen this to be true of backward integration It is true also of forward integration, for a firm which is forced to sell to a monopoly will be likely to try to develop an alternative market by engaging in the later stages of manufacture on its own account. Recent illustrations of integration from fear of monopoly are the erection of steel furnaces by Morris Motors, Ltd. (backward integration), and the efforts of British film producers to operate chains of cinemas in opposition to American-owned circuits (forward integration). There is another way in which monopoly may give rise to vertical integration. If powers of monopoly are exercised jointly by a number of "independent" firms, which share the market either through custom, or agreement, or a cartel, it may be difficult for the more efficient of these firms to expand horizontally by increasing their share of the market. Vertical integration then provides a loophole for expansion—first, because horizontal can be

[1] H. Levy: "The New Industrial System," pages 118-122.

deflected into vertical expansion; and, second, because output which is worked up in later stages inside the firm is subject to less rigorous control (or to none at all). It is generally easier for a firm to obtain an increase in its quota if it is prepared to keep the increase "off the market" by creating a new market in its own works.[1]

Vertical integration is not only a reaction to monopoly, but very often a method of extending monopoly from one stage of production to another. In many industries there is some bottleneck, control of which carries with it control over the whole industry. Strategic importance may attach, for example, to mineral deposits, or to transport facilities, or to the channels of distribution. If any of these come to be monopolised, the monopolist can bring pressure to bear on producers at other stages, and can enter into competition with them so as to extend its monopoly advantage. The Standard Oil companies, for example, used to have a monopoly of pipe-lines for the transportation of crude oil from some of the American oilfields. Their refineries were given an advantage over independent refineries by the charging of transportation rates in excess of costs, and by onerous shipping requirements such as high minimum shipments of oil. The companies' advantage in refining was reinforced by their monopoly of the tank-wagon wholesale delivery of oil—a service which was on too large a scale to be readily undertaken by a rival company and which other companies were not free to use. Similarly, the ownership of refrigerator cars, and stockyards, by the large American meat-packers helped to preserve them from competition in the main business of meat-packing. The Aluminium Company, having a monopoly of American deposits of bauxite (the ore from which aluminium is made), has completely eliminated competition in the fabrication of aluminium products by maintaining a high price for ingot aluminium.[2]

The extension of a monopoly advantage by vertical integration is only worth while if the added stages of production more than pay their way; there must be a prospect either of economies of scale or of real monopoly gains. If the monopolist can make no more profit than the firms which he ousts from the market, he may have cause to regret his hastiness. In practice, he may be expected to do better than his competitors, for he has power to capture a large market by driving them out of business, and if the added stages lend themselves to large-scale production, his costs will be lower than theirs.

The trend towards vertical integration is much less powerful than the trend towards horizontal integration. It is obstructed by lack of familiarity with the technique of other stages of production and

[1] In the long run, of course, other producers feel the competition of the increased output just as acutely as if it had been sold in the open market.

[2] cf. A. R. Burns: "The Decline of Competition," pages 439-443.

by lack of capital with which to finance vertical expansion.[1] Above all, it is obstructed by the forces which make for the specialisation of firms on a narrow range of processes. Disintegration goes on side by side with integration as one industrial process after another reaches the scale at which it can be turned over to specialist firms. The mediæval scribe is displaced by firms of publishers, printers, paper makers, machine makers, and a host of others.[2] The livestock industry breaks up into cattle breeding, cattle rearing, cattle fattening and the preparation of feeding-stuffs with all its ramifications. Even in the steel industry, tin-plate manufacture, tube making, heavy forging, and many other branches still resist integration.

(c) **Lateral Integration.**—Lateral integration is the turning out of additional products or styles of product. When, for example, a railway runs refreshment rooms, hotels, steamships, and so on, it is providing services which are connected *laterally* with rail transport; whereas, if it builds its own locomotives it is expanding *vertically*, since locomotives are essential to rail transport, while refreshment rooms, fortunately, are not. In this example, lateral extension is undertaken in order to supply markets with which a connection has already been built up—railway passengers need refreshment, hotel accommodation, etc., and it is convenient and economical to have them provided by the railway. In addition, there is a spreading of risk, especially if a *competitive* service like road transport is integrated; total receipts from all the company's activities will be more stable than receipts from rail transport alone.[3] The use of the same raw material in a variety of products, or the fact that the same technical problems are involved, may also lead to vertical integration. A firm which produces motor-cars, for example, may branch into aeroplane manufacture. Again, firms frequently take up the manufacture of new products which have been discovered by their research department or which have been brought to them by inventors unable to obtain

[1] cf. Marshall: "Industry and Trade," page 216: "A firm with limited capital can seldom undertake considerable vertical expansions with success; for such expansions are not easily made by gradual steps. On the other hand, a business may proceed gradually and tentatively when extending its operations horizontally in the same stage." (Quoted by Marquand, *op. cit.*, page 44).

[2] cf. Allyn Young: "Increasing Returns and Economic Progress," *Economic Journal*, 1928.

[3] Integration of competitive services also introduces the monopoly motive, since control (or suppression) of substitutes is the basis of monopoly power. Railways may engage in road transport in order to protect themselves against competition from independent road transport companies. Lateral integration may also be a method, not of defending but of exploiting a monopoly advantage. A firm with a monopoly of one important product may put pressure on dealers to handle its entire line of products ("full-line forcing") by threatening to deprive them of its monopolised product. By extending its line of products, it can take full advantage of its bargaining strength.

a backing elsewhere. Sometimes a new department or branch factory is added, but more commonly a subsidiary company is formed to manufacture the new product. Lateral integration is also a line of retreat or of advance when competitive pressure changes. If, for instance, competition develops in the poorer qualities of a product a firm is likely to fall back on the special qualities in which its reputation and connections help to insulate it from competition; in falling back, it is likely to look for new styles and brands which will compensate for its loss of trade in standard lines. A firm which is expanding may vary its output for similar reasons; once it begins to bump up against the limits of its market it may find it easier to make headway by branching out into new lines rather than by pushing on with the old ones.

Except in industries where the products are highly standardised, lateral growth is almost as common as horizontal. Indeed, if we put aside combination and the setting up of branch factories (both of which are predominantly horizontal), and think in terms of the establishment only, we shall probably find that change from one line of production to another plays at least as large a part in growth as horizontal expansion and contraction.[1] This does not mean, however, that firms are tending to make a wider variety of products. Lateral integration as a process of change goes on all the time, but lateral integration as a state of affairs is probably making little progress. Firms switch from one line to another, but do not necessarily add new lines to old.

**(d) Territorial Integration.**—The growth of the firm is frequently accompanied by geographical diffusion—the planting of branches over a wider area, or union with firms in other parts of the country. In some industries, indeed, diffusion and growth are almost indistinguishable. Chain stores, railways, electric supply companies, and other firms engaged in transport and distribution, find extension over a wider area much the easiest path of growth. In other industries diffusion is less attractive. The tremendous reduction in transport costs which has taken place over the last century has made it possible to concentrate industry in large central units which supply an area formerly served by a multitude of small local firms. The economies of large-scale production are often more than enough to compensate for the extra cost of drawing on a large area of supply, or of radiating output over a wide market. The big firm on a central

[1] Take, for example, the history of a sawmill started near Stockton in the middle of last century. During the war, the company turned to making shell-boxes on a large scale. In 1924 it took up joinery manufacture, supplying housing schemes at Billingham and elsewhere. Later, a licence was obtained to manufacture patent wood casement-windows, and business expanded so rapidly that new factories had to be put up. In 1935 a further factory for the manufacture of flush-panel doors was built. Now the company has begun to produce two-seater aircraft. (P.E.P. Broadsheet No. 68, 1936).

site can do things more cheaply, in spite of the handicap of a bigger outlay on transport, than small local producers scattered through each village. A familiar illustration is the decline of farm butter- and cheese-making and the concentration of dairying in creameries.

The battle between economies of scale and market resistances,[1] therefore, has gone heavily in favour of economies of scale, and by implication, it would seem, in favour of centralised production also. But the implication is incorrect. The improvement in transport and communications has made possible not only large-scale production on a central site, but also the union of scattered plants under the same management. The railway, the Post Office, and the telephone, make it easy for a firm to manage plants hundreds of miles apart—a feat which would have been almost impossible two hundred years ago. Thus where the main economies of scale are not *technical*, and can be realised in a comparatively small plant, the firm tends to grow by diffusion, not by concentration. Again, if there are strong forces making for lateral or vertical integration, and there is a locational pull on the integrated products or stages of production towards different districts, firms will be forced to operate branch factories in each of these districts. If a firm which manufactures rubber tyres wishes to own rubber plantations and textile mills it will probably have to go to Malaya for the first and to Lancashire for the second. Similarly, it will be unlikely to carry on the production of golf balls or rubber cushions in the districts in which its tyre factory is located.

## The Implications of Large-Scale Production: A Reconsideration of the Optimum Firm.

We have studied in this chapter the ways in which firms grow and the obstacles which lie across the path of growth. In earlier chapters[2] we studied rather similar problems, but from a different point of view. We took what might be called a "mechanical" view of the firm, looking at it in isolation from its industrial environment, so as to discover what size of firm was most efficient. We started out from the conception of an optimum firm, which would be able to produce at less cost than firms larger or smaller in size; and we were careful to distinguish between such a firm, with its costs at the lowest possible level, and firms which, because of some monopoly advantage, found it easier to earn a profit and were better fitted, therefore, to survive competition. The optimum firm, in short, was an economic norm against which we could measure existing firms.

But—and here we come to the link with the present chapter—we cannot take a consistently mechanical view of the firm, abstracting from the industrial environment, and juggling with different sizes

[1] See above, page 80.

[2] See Chapters 5 and 6.

of firm until we find the optimum. Firms exist in association and in competition with other firms; they are linked organically with one another.[1] The optimum firm, therefore, must be looked at *biologically* against a background of market conditions from which we cannot abstract. Its costs depend not simply on how it does things (that is, on what is happening inside the firm), but also on what it has to do, and what it has to do depends upon the industrial environment. If the firms in an industry are of one size they find themselves faced with one set of problems; if they all double in size they will be faced with quite a different set of problems. But the firm which would be most efficient in coping with the first set of problems is very unlikely to be of the same size as the firm which would be most efficient in coping with the second set. The size of the optimum firm depends upon the particular organisation of industry into which it has to fit.

Thus the growth of a firm must be considered biologically, not mechanically. What is optimum, and what sub-optimum, depends on the industrial environment within which the firm functions. If the environment changes the optimum changes, and the growth of the firm of itself may be sufficient to alter the environment.

This elementary proposition is frequently lost sight of in arguments for and against nationalisation. It is suggested, for example, that nationalisation will result in the creation of excessively large firms, since it will leave only one firm, run by the State, even in industries where the optimum is demonstrably small. But is this really so? Is it not possible that nationalisation may of itself create an environment favourable to much larger units than can survive free competition?

In the first place, many of the risks which keep the optimum firm comparatively small are themselves the product of competition. The individual producer must adapt his plans to the plans of his competitors without knowledge of their intentions. He is in the dark about the new models and new products which they mean to bring out, the output which they propose to market, and the methods by which they hope to market that output. Thus he is kept busy meeting their competition, retaliating in kind for invasions of his

[1] Compare A. N. Whitehead: "Science and the Modern World," page 238.

"The trees in a Brazilian forest depend upon the association of various species of organisms, each of which is mutually dependent on the other species. A single tree by itself is dependent upon all the adverse chances of shifting circumstances. The wind stunts it: the variations in temperature check its foliage: the rains denude its soil: its leaves are blown away and are lost for the purpose of fertilisation. . . . But in nature the normal way in which trees flourish is by their association in a forest. Each tree may lose something of its individual perfection of growth, but they mutually assist each other in preserving the conditions for survival. The soil is preserved and shaded; and the microbes necessary for its fertility are neither scorched, nor frozen, nor washed away. . . . A species of microbes which kills the forest, also exterminates itself."

market, scheming to steal a march on them by some new agreement with distributors or some new advertising campaign, giving uncertainty for uncertainty. If all this competitive effort were inevitable, the optimum firm would presumably be fairly small, since a large firm would lack the adaptability and thrustfulness necessary in a highly competitive market. But competitive secrecy and uncertainty are by no means inevitable. A monopoly gets rid of both. So also would nationalisation—the formation of a State monopoly.

Nationalisation would also eliminate risks due to fluctuations in trade and employment. The individual producer has always to run the risk of recurrent slumps in trade—slumps which he can do nothing to prevent. He cannot expect to operate his plant continuously at capacity and will be likely, therefore, to avoid equipment which, however efficient, is worth while only if in continuous use. If all industry were under the control of the State, however, it would be possible to regulate output in accordance with a comprehensive plan. The risk of discontinuous operation would be greatly reduced and larger establishments would become economical. At the same time, many of the managerial difficulties which, in a fluctuating market, keep the optimum low would disappear immediately producers were offered a stable market. The task of management would be to supervise the production of an output fixed in advance, while price—and sales—policy would be controlled from above, much as in the cartel.

There is a third way in which, as a result of nationalisation, the industrial environment may become more favourable to large-scale production. Competitive pressure often induces a firm to make a wide variety of products, and this variety, by complicating the problem of management, tends to limit the growth of the firm. In a socialist community, however, it would be possible to carry specialisation a good deal further, and to concentrate particular products on particular firms. This would allow each product to be produced on a larger scale than before without giving rise to managerial diseconomies.

We cannot assume, therefore, that there is necessarily a loss in efficiency when an industry is taken over by the State and run on a scale greater than the competitive optimum. State control and private enterprise are different forms of industrial organisation, and the same level of administrative efficiency within the two forms of organisation may yield quite different results. Nationalisation creates opportunities of economy that may compensate for the diseconomies of large-scale organisation ; it also changes the temper of management and sets up a more elaborate hierarchy with a powerful bias in favour of precedent and prudence.

The issues raised by nationalisation are much more far-reaching than can be discussed at this stage ; we shall return to them at more length in Chapter 29.

# PART III—SUPPLY AND DEMAND

## CHAPTER 10

## THE EXISTING ORDER: A PRELIMINARY SURVEY

WE speak of a Capitalist "System." But is there a system at all? Are there not constant gluts and shortages to prove that too much of this and too little of that have been produced? Or chronic unemployment to prove that our resources are often not used at all? There is no comprehensive scheme by which, for example, the citizens of London are fed, clothed, and housed. No one arranges for the supply from day to day of the exact number of gallons of milk, or loaves of bread, or pounds of sausages which are daily consumed. Probably no one knows. No one arranges that priority in the use of resources will be given to industries of first importance. We are free to go on producing what no one wants or what is over-abundant, however pressing the need for our services in some other job. No one arranges that sheep will provide mutton and wool in just the combination that satisfies the demand for each. Farmers can choose their own breed of sheep in complete ignorance of what other farmers are doing, where their mutton goes to, or who is clothed in their wool.

So much freedom of choice, combined with so much ignorance, might seem sure to lead to chaos. Yet the citizens of London are fed, the industries of first importance do not disappear, and wool and mutton find buyers without too much difficulty. Resources are mobilised to meet our wants with *some* appearance of order. Is there, then, some "invisible hand" steering resources into the right channels just as large-scale planning might do? Is there, after all, a system?

We must be careful, first of all, not to under-rate the amount of planning which is already undertaken by individual firms. Each firm has to decide what articles to produce, where to produce them, how much of each to produce, and what methods of production to employ. A large firm, like Imperial Chemical Industries, or Stewart and Lloyd's, is faced with broad questions of policy that are not very different, even in scale, from the problems that a State Planning Authority would have to solve. In smaller firms planning is narrower in scope. But there is a daily task of organisation, of planning what is to be done, and giving instructions who is to do it. This planning of details can be undertaken by firms which have little knowledge of what is going on elsewhere. It is often unsatisfactory because there is no one

whose business it is to co-ordinate the plans of each firm. But the work of organisation undertaken inside the firm, however unsatisfactory at times, does represent a contribution to planning which would have to be made somehow if universal planning were introduced.

Secondly, we must remember that consumers do their planning too. Each consumer plans his own expenditure, buying first what he most requires and spending what is left of his income to satisfy his wants in decreasing order of urgency. If what he wants greatly is scarce, he offers a high price and so induces someone to supply him with it. On the other hand, he penalises anyone who persists in trying to market too much of a commodity (in relation to the supply of other commodities) by paying only a low price for it. Now prices govern profits, and the prospect of profit shapes the plans of producers, drawing them into industries where they are good, and drawing them out of industries where they are poor. At the same time, the tendency for consumers to buy in the cheapest markets, and for the factors of production to sell their services where the highest price is offered, drives high-cost producers to the wall and reinforces the motives of the survivors to use the most efficient methods of production. Those who charge high prices cannot find a market: those who pay low wages cannot find employees. The staying-power of a firm or industry —in the absence of State assistance[1]—depends upon its ability to show a profit; and this profit—if competition is working freely[1]—is a sign that consumers have a more urgent demand for the product of the firm or industry than for the amount of any other commodity that could be produced at equal cost.

Not only, therefore, is there a *system* of rewards and punishments—the spur of profit and the penalty of loss—knitting together the plans of consumers and producers; it might also seem as if the system were an *ideal* one—that the incentive of private profit works to secure just those results which are most in the public interest, and that large-scale planning, or indeed government intervention of any kind, is superfluous. It was this rosy view of the workings of supply and demand that crystallised in the doctrine of laisser-faire. In the twentieth century, with the hard facts of unemployment, poverty, monopoly, and a hundred other evils staring us in the face, supply and demand, and laisser-faire are fallen idols. Private profit and social advantage—as we shall see over and over again—do not coincide. The system is far from ideal: but the main point, for the present, is that there *is* some system and not just chaos.

Capitalism then may fairly be spoken of as a system. But what is it that distinguishes it from other economic systems? An essential feature of Capitalism is the private ownership of property. In capitalist society men have the right to accumulate property without limit for their exclusive use, and the right to dispose of it as they choose so long as they keep within the law. Power of disposal over property carries with it as a corollary a second feature of Capitalism—free

[1] These are very big qualifications. More will be said about them later.

enterprise. Capitalists have the right to sink their capital in any business, subject to various legal restrictions. If they see a way of improving on existing methods of production, they are free—except in one or two industries where there are legal obstacles to competition—to back the improvement, at the risk of losing their capital, by entering into competition with established firms. Or if—to take a less commendable example—they see a way of overcharging the public by making a corner in some "key" commodity, they are free to put up the necessary capital and take what profit they can.

More important than the enjoyment of these rights by individual citizens is the inequality with which they are distributed. Ownership and control of property are practically the monopoly of a single class. Five-sixths of the National Wealth—or rather, of that part of it which is in private hands—is owned by the small class of persons (6 per cent. of those who are over the age of 25) who have each at least £1,000; three-fifths is in the hands of the still smaller class (1 per cent. of those who are over the age of 25) who have at least £10,000.[1] Some slight redistribution has, it is true, taken place over the past thirty years. But it is plain that, in spite of high death duties and a steeply graduated income-tax, there has been no redistribution on the scale popularly imagined. If the rich have lost ground, it is not because their own fortunes are smaller or less valuable, but because the small capitalist with a house, an insurance policy, and a few investments, is assuming more importance.

Concentration of control has gone much further than concentration of ownership. Capitalists generally delegate power to dictate what use shall be made of their property to company directors, who may own no more than a tiny fraction of the vast resources under their control.[2] A few hundred persons, most of them in personal contact with one another, come to exercise effective (but not unlimited) control over the nation's industries. Nominally they are the agents of their shareholders, the legal owners; but in practice, shareholders, even when (as rarely happens) they are well informed and organised, are generally powerless to resist the decisions of their "agents."

Here a qualification must be made. Rights of ownership are not always effective: but neither are powers of control. If owners delegate their powers to company directors, they in turn must delegate responsibility to various underlings. They cannot possibly decide everything. Like the owners, too, they may find their powers very limited even when they wish to exercise them. They may have been kept in ignorance, or find themselves in a minority, or be bullied by their employees or by a business associate or "controlling interest." It is almost as difficult to decide who has effective power in a firm as to decide who has effective power in the State.

---

[1] "The Distribution of the National Capital," Daniels and Campion (estimates revised by Mr. Campion).

[2] See above, page 101.

Concentration of ownership and control, however qualified, are characteristic of Capitalism. Society is divided into capitalists and workers, and into masters and men. These class divisions, as we have seen, do not altogether coincide. The manager of a business need not be a capitalist; there are salaried-managers as well as owner-managers. There are divisions, too, within each class. Capitalists may be bondholders or shareholders; some workers enjoy more responsibility or require more skill than others. Finally, class divisions are not absolute. Wage-earners may own some capital, and capitalists may draw a salary. Employer and employed, capitalist and worker, can change places. Classes are not castes.

Nevertheless it is impossible to disguise the fundamental line of division in capitalist society between those who possess capital and those who do not. This is a line which it is always hard for a wage-earner to cross. True, there is an increasing number of wage-earners who are also property-owners, and who supplement their wage-income out of interest and dividend-payments. But their income from interest and profit remains a negligible fraction (at most 5 per cent.) of the total income of all wage-earners. It is true, too, that there is an increasing number of salaried workers who, though in a sense wage-earners, have accumulated some capital and are generally "capitalist" in outlook and sympathy. They are not, however, *numerically* of such importance as to upset the broad classification of society into wage-earners dependent for their livelihood on selling their labour, and capitalists drawing a large part of their income from rent, interest, or profit. Wage-earners, independent workers, and salary-earners with less than £250 per annum, form altogether 90 per cent. of the occupied population. But their share of the National Income is not much greater than that of the remaining tenth. Of the income of the upper tenth, perhaps half comes from rent, interest, and profits, and half from professional and managerial work. Thus, while there may be no clear line of division between those with capital and those with little or none, there is no blinking the fact that for some people income comes mainly from work while for others an important, sometimes a major, contribution comes from property.

This class division is a fairly rigid one. Some wage-earners contrive to amass fortunes and some capitalists to squander them. But for the vast majority, poverty breeds poverty and money makes money. The upbringing, the education, the social contacts, the inherited wealth, of the son of a property-owner give him a start in life for which even outstanding ability in a wage-earner may never make up. The one may be trained by abundance of opportunity to habits of leadership which, in the world of business, command a price commensurate with their rarity. The other, lacking in influence, or drive, or knowledgeability, or capital, cannot strike out on his own, but has to work his way up patiently in a firm which may be very slow to recognise his ability. Capital can be left to accumulate at compound interest; but labour can set nothing aside until it is fed, clothed, and housed. If a worker is poorly paid, he cannot keep in a proper state

of efficiency; his value in the labour market falls, and he may cease to be worth even his poor wage. Thus it is hard for the very poor not to sink lower, and for the very rich (in the absence of high taxation) not to grow still richer. For the rest of the community, the levelling influence of education and the growing importance of capital in industry roughly offset one another.[1]

The same class division is apparent if we look at Capitalism from another angle. In any economic system there are three main functions that must be undertaken if the work of specialist producers is to be co-ordinated. There must be management: some method of planning the tasks of workers who co-operate in a firm. There must be investment: some method by which the firm procures its capital. And finally there must be risk-taking: willingness to face the prospect of loss in the hope of gain. It is the peculiarity of Capitalism, not that these functions are necessary (they are, in fact, inescapable in any industrial systems) but that each of them is undertaken by a definite class of persons. Inside the firm, instead of everyone having some control over planning—either by direct participation or by electing their overseers—decisions are taken by "managers" or "owners" who strike a bargain with their employees but are not otherwise responsible to them. The capital of the firm is drawn, not from its employees, or from the mass of workers, but from the small class of investors who own the vast bulk of the country's wealth. And the risks of the market are taken, not by wage-earners working under a contract of service, but by a class of risk-bearing specialists, who allow their capital to be used without the certainty that it will be returned intact, and in the hope that it will be returned with profit. The wage-earner is paid an assured sum in return for work which may involve his employer in loss: he transfers the risk of loss in his work to his employer.[2]

Thus each of the three functions is undertaken by a limited class of persons who are not directly responsible to the workers whose efforts they co-ordinate. The work of co-ordination rests with "private enterprise"—i.e., with capitalists and their agents. But it does not rest *entirely* with private enterprise. There is, first, the Co-operative Movement. A co-operative society resembles a joint-stock company in that the management is in the hands of a salaried official appointed by the shareholders or their nominees. But it

[1] Even in the United States the chances of a wage-earner rising to the master class are very slight. Only one in ten of "business leaders" is the son of a wage-earner. On the other hand, one in two of all "business leaders" in the United States is owner or executive in what was once his father's business. Nor are the wage-earner's chances of promotion improving, for it is from the sons of business men that recruitment is increasing. ("American Business Leaders," F. W. Taussig and C. S. Joslyn).

[2] The burden of other, and graver, risks (of ill-health, sudden death, unemployment, etc.) is not transferred. The Health, Accident, and Unemployment Insurance Schemes, however, have now relieved the worker of part of the burden.

differs in that the shareholders who supply the capital are also the customers who buy the product. There is thus no conflict of interest between profit-seeking shareholders and cheapness-seeking consumers. The work of co-ordination is simplified, while the consumer gains through the growth in power of bodies devoted exclusively to his interests.

Secondly, there is the State—the co-ordinator of co-ordinators. The State may pass Acts limiting freedom of enterprise and prescribing the terms on which business of any kind must be carried on (e.g., in the Factory, Truck, and Marketing Acts). It may set up bodies like the Central Electricity Board, the Port of London Authority, or the B.B.C. to supersede private enterprise. It may subsidise or nationalise if the machinery of the market breaks down or individualistic planning leads business men into a tangle. It may give Trade Boards power to fix minimum wages and maximum hours if workers complain of exploitation. Thus the State is not simply an umpire interpreting the rules of the game. It is umpire, M.C.C., and cricket team all in one. It makes the rules and enforces them; and it has at least one team of its own. But it must not have all the teams (or even a very large proportion), or we have left Capitalism behind and adopted a Socialist régime.

## CHAPTER 11

## PRICING

### The Function of Prices.

THE central problem of economics is the accommodation of scarcity to human want—in other words, of supply to demand. An understanding of the mechanism by which supply and demand are balanced, therefore, is fundamental in every branch of economic theory. This mechanism is generally referred to as "pricing" or "the pricing process"; for it is about prices that supply and demand pivot. If supply increases, while demand remains constant, prices will fall; if demand increases, while supply remains constant, prices will rise. These movements in prices tend to restore the balance between supply and demand. If there is an excess of supply over demand, or of demand over supply, prices tend to move so as to wipe out the excess and bring supply and demand back into line with one another. It is this sensitiveness of prices that enables them to knit together the plans of producers and consumers. They can choke off an over-supply by falling below cost of production to levels unprofitable to producers; and they can choke off an excessive demand by rising to levels at which consumers do not think it worth while to buy.

In balancing supply and demand, prices perform an important social function. On one side, they reflect our "values," our

estimates of how much things are worth; on the other, these reflect the scarcity of things, and the cost of making them available to consumers. They preserve a balance between value and cost. But the balance may not be struck in the best possible way. Prices may reflect our wants imperfectly (for example, because of ignorance); and they may get out of line with costs (for example, because of monopoly). Market prices, that is, may diverge from those ideal prices which would accurately reflect our wants and the cost of meeting them. For some commodities the price will be too high and for others too low.

From a slightly different angle, the function of price is to ration a limited supply of goods amongst consumers. The highest bidders, for whom the goods have great utility, are able to buy what they want so intensely, while those who think the price too high are forced to content themselves with other goods which they consider a better bargain. Similarly the price paid for land, labour and capital rations the limited supply of productive resources between competing industries. Those industries which can offer the highest prices for the factors of production are able to commandeer supplies; while industries which find these prices (rent, wages or interest) too high to yield a profit are forced to contract, surrendering the resources in their employment to better-placed industries whose goods are in more urgent demand. Prices serve to distribute goods and resources to the points at which they are scarcest in relation to the wants of consumers. Accurate pricing is the essence of economy.

But pricing is not in fact accurate. Goods fall to those who are prepared to pay the highest prices; but high prices mean less to a rich man than to a poor man, so that those who buy goods are not necessarily those for whom the goods have most utility. A poor man might be willing to make great sacrifices for the sake of a visit to London, a motor-car, a university education for his son, and so on; a rich man might attach comparatively little value to any of these things. But he might pay for them none the less, while the poor man found them beyond his means. Similarly, if consumers are hasty or ill-informed in making purchases, the goods which are produced will not be those which are of most value in meeting their wants. Finally, the output of some goods may be unduly limited through monopolistic control. Pricing, therefore, may fall short of the ideal. Too much of one commodity, too little of another, may be produced. Of our limited productive resources an excessive share may be put at the disposal of one industry and a deficient share left for another.

**The Theory of Price or Value.**

We are faced, therefore, with two distinct sets of problems. First, we have to analyse how pricing does in fact work; and, second, we have to analyse how, ideally, it ought to work in the best interests of society. The first problem is one of explaining how the prices of different commodities come to be what they are. We have to ask questions like: Why can we buy several tons of coal for the

price of a fur coat? Why does a first-class detective story fetch only 6d. when dull treatises are valued at 30s. and upwards? Why are some prices above, and some below, cost of production? What would happen if the Government attempted to fix the price of bread? The second problem discussed in the theory of price, or value, is that of what prices should be. This problem may be approached from two different angles: the angle of the just price and the angle of ideal output. We may ask: on what principles are we to decide when a given price is just and when extortionate? Or we may ask: on what principles of price-fixing will we secure an optimum allocation of productive resources between alternative uses? In short: how ought we to value the products of one industry as compared with the products of others?

It should be observed that the second problem lends itself to scientific treatment much less readily than the first. When an economist discusses how prices *are* determined in the real world, he will be wrong only if his reasoning is faulty, or if his pronouncements are based on inadequate or inaccurate observation. He will be wrong, that is, only if he is slovenly and incompetent. But when he discusses how prices *should* be fixed, or what principles *should* govern the allocation of resources between industries, he may find himself in disagreement with his colleagues, however careful and competent he is. For he will be raising questions of social justice on which no economist can claim to speak with authority. He will be forced to base his conclusions on his private conception of what is fair and just, and this conception may be one that is not generally shared.

**The Importance of Social Institutions.**

It is essential to study how prices are, or should be fixed, against the background of a given organisation of society, in a given context of social institutions. We can assume institutions broadly similar to those with which we are familiar—private property, inequality of income, a wage- and money-system, unemployment, joint-stock enterprise, and so on—or institutions widely different from our own: the institutions, for example, of a Socialist State, or a South Sea island, or a world of monopolies. It is natural, in a discussion of price-determination, to assume the institutions of a capitalist economy, since the questions posed and answered would otherwise have little relevance to the society in which we live. This is not to say, however, that a discussion of the forces determining prices in other types of society is superfluous or uninteresting. Such studies might be of great assistance in enabling a decision to be made on the merits of alternative sets of institutions. An understanding of the machinery of price-fixing in a Socialist State, for example, might make it easier to weigh up the advantages and disadvantages of Capitalism.

If we are discussing how prices *should* be fixed, an institutional background is equally necessary. Prices which are fair in a society governed in one way may not be fair in a society governed in another. For instance, it may be fair in a society of teetotallers to charge

high prices for wines and spirits to a small minority of hardened drinkers; while in a society of hardened drinkers such prices might rightly be considered outrageous. In a land of slum-dwellers it may be fair to make house rents specially low to facilitate slum clearance; whereas low house rents might not be at all fair in a country free from slums. Where women workers are mainly engaged in earning pin-money it may be fair to pay them less than men for work of equal value; whereas "equal pay for equal work" might be the proper maxim where men and women were equally dependent on wage-earnings for their livelihood.

This point is of particular importance when we are discussing the justice of price discrimination—that is, of charging different prices to different groups of consumers for the same service. Suppose, for instance, that some consumers are rich and some are poor, and that the distribution of income is universally admitted to be unfair. Then it may be right to charge rich consumers higher prices than poor consumers. Arrangements may be made for selling potatoes, or milk, or wireless sets, or cinema tickets to the poor below the current market price. But if people's money incomes were admittedly fair, or approximately fair, such arrangements might properly be condemned.[1] Similarly, discrimination might be introduced with a view to assisting unemployed or disabled persons when in different circumstances the right course would be to increase the incomes of these persons. Fairness and expediency cannot easily be disentangled, and expediency turns very much on the social background against which price-fixing takes place.

**The Assumption of Perfect Competition.**

We have seen that the theory of prices is an attempt to explain how prices are, and how they should be, determined. These two parts of the theory, however, are very easily confused. Economists, in analysing what does in fact govern prices, are forced to make several simplifying assumptions. They generally begin by assuming that competition is "perfect"—that is, that no individual producer or consumer has any control over market prices. This assumption is unrealistic—increasingly unrealistic because of the growth of monopoly, advertisement, etc. But it is nevertheless necessary if we are to make sense of the workings of competition. If we start from the extreme of "perfect" competition, we can go on to study the ways in which competition is limited in the real world[2]; whereas if we drop the thread of "perfect" competition we shall find it hard

[1] Of course, there might still be good grounds for price discrimination on an income-basis if it led to *economy*; for instance, where goods could only be produced cheaply if sold on a large scale and could only be sold on a large scale if discrimination were practised (see below, pp. 211-2.

[2] When economists believe that competition is *in fact* not far from perfect—once a fairly common belief—their theories can easily be interpreted as a complicated apology for the existing social system.

to make sense of the workings of competition and monopoly. Now many economists have tended to make perfect competition not simply a point of departure from which to analyse what does govern prices, but also an ideal goal towards which prices should be made to move. They have held that it is socially desirable to reproduce conditions of perfect competition as nearly as possible. This is not, as we shall see, an extravagant view. But it is just as well to keep an open mind on the subject, and to distinguish carefully between the forces which do and the forces which should determine prices, until the first part of the theory, to which we must now turn, has been fully discussed.

**Price and Value.**

The theory of prices is sometimes spoken of as the theory of value.[1] This alternative title is intended to make it clear that it is exchange values, or *relative* prices, that are under discussion. Suppose, for example, that pears cost 3d., plums 1½d., and apples 1d. Then the exchange value of pears is two plums or three apples each. Suppose now that money loses half its purchasing power and that the price of pears goes up to 6d., of plums to 3d., and of apples to 2d. The exchange value of pears in terms of plums, apples and other commodities is unaltered. All that has changed is the value of money in terms of commodities. Thus a rise in the price of plums may reflect a change in their value relatively to other commodities or relatively to money. The theory of prices, therefore, is faced not with one problem but with two. It must explain what causes a change in relative prices or exchange values; and it must explain what causes a change in the value of money. The first problem is dealt with in the theory of value, the second in the theory of money.[2]

[1] At least five different meanings of the term "value" can be distinguished: moral good, æsthetic merit, utility, exchange-value, and ideal exchange-value. When we say, for example, that we "value" freedom, we are using value in the moral sense. We think that freedom is something to which men are morally entitled, something that they should have; we are certainly not thinking in terms of supply and demand. Again, if we set a high "value" on Shakespeare's plays or Beethoven's music our judgment is an æsthetic one. If we speak of water as being "valuable," we mean that it answers to some well-known human needs, not that it fetches a high price; we are thinking of its utility. The "value" of a house may refer to its selling or exchange-value or to the price which some imaginary purchaser might or should be willing to pay for it (for instance, we speak of house-property fetching less than its "value.") In economics "value" is sometimes used to mean exchange-value, and sometimes to mean utility, or "value-in-use."

[2] A strict separation of the two problems is not possible. A change in relative prices, for example, may be sufficient to precipitate a change in the value of money; and a change in the value of money is invariably accompanied by changes in relative prices. In elementary economics, however, it is legitimate to assume that the two problems can be disentangled: otherwise no progress towards understanding either is ever likely to be made.

**Early Theories of Value.**

Early theories of value were theories of what governs prices *in the long run.* Day-to-day fluctuations in price—fluctuations in "the temporary or market value" of a commodity—were put down to "the higgling of the market," or dismissed with a vague reference to supply and demand. Analysis was confined to "normal" or "natural" values—that is, to the long-run trend of prices, about which day-to-day fluctuations take place.

**The Labour Theory of Value.**

Most early theories were variants of the Labour Theory of Value. According to this theory the "natural" value of a commodity depends on the amount of labour embodied in it—on "the toil and trouble" bestowed on its production. The value of all commodities, it was argued, is reckoned in terms of money. Since the value of money is constantly changing, however, money cannot be the ultimate standard of value. But there must be *some* ultimate and invariable standard, some circumstance common to all commodities, by which exchange values are determined. Now we know that labour is needed in the production of almost everything of value, and that commodities "are counted dear or cheap according as they can be had with much or little labour." The effort expended in an hour's labour, moreover, will always produce the same quantity of goods; labour is a reliable measuring rod, whereas money is not. Finally, it is clear that in a primitive community with little capital, exchange values will correspond fairly closely with the amount of labour embodied in each commodity. The laws of nature, which were supposed to apply in such a community, were often appealed to as a standard in the eighteenth century, and coloured much of the political and economic theory of the period. It seemed to be in accordance with these natural laws that prices should correspond to the cost, in human effort, of producing commodities.

The most striking deficiency of this theory is that it neglects the influence of *demand* on value. It is obvious that commodities acquire value just as much because they satisfy our wants (i.e., possess utility) as because they cost some effort to produce (i.e., embody labour). But while it is easy to see the relationship between effort and value it is not easy to see the exact connection between utility and value. The things that have most utility (e.g., water) are often the least valuable; while the things that are of doubtful utility (e.g., diamonds) are often of most value. It was this puzzling "paradox of value" that led Adam Smith and other economists to give up the attempt to relate utility to value and to fall back instead on labour.

The second main deficiency is that the theory gives only a partial explanation of the influence on value of *supply.* It is plain that, of a number of things in equally urgent demand, those are most valuable that are most limited in supply (i.e., that are scarcest). But it is not so plain that those things are scarcest, and consequently most

valuable, that require most labour to produce. *One* reason why things are scarce and valuable is that human labour is needed in order to produce them. But there are other reasons. There is, for example, the initial scarcity of land and raw materials with which labour co-operates. If a ton of tobacco sells for more than a ton of wheat, the explanation may lie just as much in the comparative scarcity of land suitable for tobacco-growing as in the comparative laboriousness of tobacco-growing. Another factor governing the relative supply of commodities, and hence their exchange values, is the scarcity of capital. Dress-making, for example, uses far more labour in proportion to capital than electricity-making. If capital becomes more abundant, therefore, it is reasonable to expect a fall in the value of electricity in comparison with the value of gowns, even although no change takes place in the labour required to make a gown or a unit of electricity. Finally, it is often the commodities that take the largest number of labour-hours to produce, or call for the most intense effort, that are cheapest. The products of the navvy fetch far less than the products of the lawyer from the point of view either of duration or intensity of toil. The reason is that navvies are abundant while lawyers are scarce. In short, it is not labour, but scarcity—or supply, which is just scarcity the other way round—that governs value—jointly, of course, with demand. Labour—toil or effort—is one thing among several that helps to moderate scarcity and so to create value.

Sometimes the labour theory of value is supported by an appeal to justice and the rights of man. It is suggested (for example, in the Marxian version of the theory) that labour alone creates value, and that, if market values are in excess of wage costs, then the labourer is being exploited and is receiving less than the value of his product. The surplus of rent, interest and profit falls to the capitalist class, not in return for any service which they perform, but because they happen to be the owners of the instruments of production. This is not so much a theory of value (i.e., of relative prices) as a theory of distribution (i.e., of the earnings of the factors of production). It is unnecessary, therefore, to discuss it in detail at this stage. As we shall see, the assumption in the theory that labour alone creates value is false, even if "labour" is interpreted broadly to include skill and judgment. The activities of saving and risk-taking, for example, can be just as productive as labouring; they are almost as essential to the creation of value.[1]

**The Cost of Production Theory.**

A more modern version of the labour theory is that prices are governed, in the long run, and in the absence of monopoly, by the cost of production. This seems plausible enough, and is in harmony with much that goes on in the business world. Nevertheless the theory breaks down and for much the same reasons as the labour

[1] See below pp. 262-5, 272-4.

theory breaks down. First, it neglects the influence of demand. We have already seen that costs of production generally vary with the scale of output—that is, that there is no single cost of production for each commodity irrespective of the amount of it that is being produced. Now this amount is presumably equal to the amount that consumers want to buy, i.e., to the *demand* for the commodity. Until we know what the demand is, therefore, we do not know the scale of output of the commodity, and we also do not know its cost of production. But if we do not know the cost of production until we know the demand, it is useless to tell us that price is governed by cost. Demand (or utility) and supply (or cost) are like the two blades of a pair of scissors (to use Marshall's metaphor)—it is as idle to argue that one or other governs prices as to argue that it is this blade or that that does the cutting.

Secondly, the theory is emasculated by the qualifications "in the absence of monopoly" and "in the long run." In modern business, it is monopoly that is the rule, free competition the exception. There are elements of monopoly in every trade, and it is these elements of monopoly that raise the most difficult problems in economic analysis. Such an analysis cannot proceed solely in terms of cost of production; there is nothing to compel a monopolist to fix his prices at or near to cost of production. We must bring in demand if we are to develop a theory of prices under monopoly. Similarly, it is not denied that prices may be above or below cost in the short run; the theory refers only to what takes place in the long run, given time for changes to work themselves out. But the short run is much more interesting than the long run; as a famous economist has reminded us, in the long run we are all dead. And we cannot say how far above cost, or how far below, prices will range in the short run until we have made a study of demand.

Thirdly, the theory gives only an incomplete account of the influence of supply. Cost of production seems simple and unambiguous enough. But in fact there are few phrases in economics so hard to interpret. Firms vary enormously in efficiency: some go on producing, year after year, at high cost, while others have costs consistently lower. In which firms are we to measure cost of production? Or again, when a firm produces several different commodities in the same building, how are we to distribute the overhead costs of the building between each of the different commodities? If sheep *will* persist in providing us with mutton and wool simultaneously, how are we to distinguish the cost of wool from the cost of mutton? To these questions, the cost of production theory returns no answer.

The deficiencies of the cost of production theory will become clearer when the modern "marginal" theory of value has been explained. Before we turn to this theory, however, we must clear away some of the confusion that clings to the terms "supply" and "demand."

# CHAPTER 12

## SUPPLY AND DEMAND

THE price of an article depends upon two distinct sets of forces, generally referred to as supply and demand. "Supply" means the quantity offered for sale by producers, and "demand" the quantity that consumers are willing to buy. Thus the "supply" of coal does not mean the amount of coal lying underground waiting to be mined (except, of course, in other contexts); it means the amount of coal which coalowners are willing to put on the market at some specified price. Similarly the "demand" for coal does not mean the amount of coal which people need, or would like to have, but the *effective* demand, the amount which people are willing to buy at some specified price. It should be observed that we cannot speak of demand or of supply without specifying some price. There is, for example, no such thing as *the* demand for coal. People will buy more coal when it is cheap than when it is dear. Similarly with supply. Coalowners will be more willing to mine and sell coal when the price is high than when it is low.

Thus demand and supply vary with price. But they also vary with other things—for instance, with people's tastes, or the standard of living, or the state of technical knowledge. The changes that take place in demand and supply in response to changes in these other things (what are called "demand *conditions*" and "supply *conditions*") differ, however, from the changes that take place in response to changes in price. Demand conditions and supply conditions are generally independent of one another, whereas price affects demand and supply simultaneously. A change in the technique of coal-mining will affect the supply of coal, for instance by lowering the cost of mining it, but it will be unlikely to react directly on the demand for coal. A change in the price of coal, on the other hand, will affect both the supply of coal and the demand for it.

The fact that supply and demand both respond to changes in price means that supply and demand can be balanced if an appropriate price is charged. The higher the price, the more will be supplied and the less demanded; the lower the price, the less will be supplied and the more demanded. Any gap between supply and demand can be closed, therefore, by raising or lowering the price. If coal-owners offer for sale more than the market will absorb at the current price, the price of coal will be forced down until the surplus disappears, either through consumers buying more or because coal-owners reduce their output. Similarly if consumers insist on having more coal, it will be through the offer of higher prices that they induce coal-owners to undertake the mining of an increased supply. Given the state of demand (demand conditions) and the state of supply (supply conditions) there will be one price (the equilibrium price) at which demand and supply can be made to balance, and competition will drive the price to the balancing point.

**Supply and Demand Curves.**

Thus, on the one hand, supply and demand depend upon, and vary with, price; and, on the other, price depends upon, and varies with, supply and demand. The inter-relationship between supply, demand and price can be illustrated by drawing a supply curve and a demand curve. These curves show how much of a commodity will be offered for sale at any price, and how much of it consumers will be likely to buy at any price. The supply curve will normally be sloping upwards from left to right, and the demand curve will slope downwards from left to right. Where the two curves intersect, demand and supply will be equal. The price at which this equality is achieved is known as the equilibrium price.

Suppose, for instance, that we are discussing the price of oranges. We can draw up schedules like those in Table I to show how many

TABLE I.

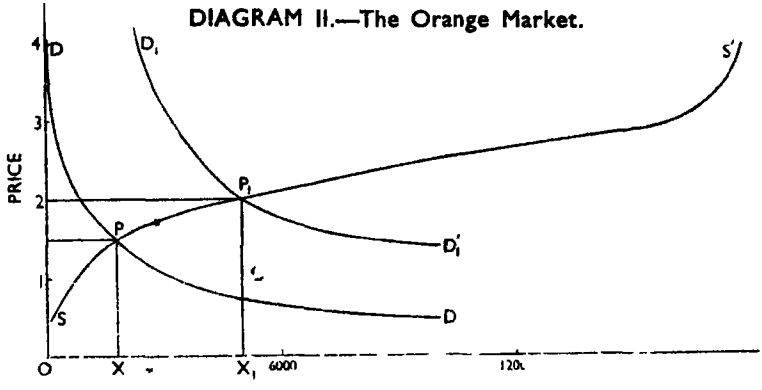
| Price of oranges | Number of oranges offered for sale (Millions) | Number of oranges consumers would buy (Millions) | Purchases of working-class consumers (Millions) | Purchases of middle-class consumers (Millions) | Purchases of well-to-do consumers (Millions) |
|---|---|---|---|---|---|
| ½d. | 50 | 10,000 | 7,500 | 2,250 | 250 |
| 1d. | 400 | 3,000 | 2,000 | 800 | 200 |
| 1½d. | 1,800 | 1,800 | 1,000 | 600 | 200 |
| 2d. | 5,000 | 850 | 250 | 400 | 200 |
| 2½d. | 10,000 | 400 | 50 | 200 | 150 |
| 3d. | 16,000 | 260 | 10 | 150 | 100 |
| 3½d. | 17,500 | 100 | — | 50 | 50 |
| 4d. | 18,000 | 15 | — | 5 | 10 |

oranges will be offered for sale, and how many will be bought, at any given price. As the price of oranges goes up, the number offered for sale goes up too; at first, very rapidly, because it will pay to stop growing other kinds of fruit and grow oranges instead; later, increasingly slowly, because there is a limit to the area on the surface of the earth where oranges can be grown. On the other hand, as the price goes up, the number of oranges bought begins to fall off. Working-class consumers are the first to be squeezed out; prices which the middle class can afford without much difficulty are beyond

their limited means. Then middle-class consumers in turn begin to drop out, while even the rich cut down their purchases until, at a price of say, 4d., demand has dwindled to insignificant proportions.

The schedules of supply and demand given in Table I, when plotted on a graph (Diagram II), give us the supply curve and demand curve for oranges. SS′ (the supply curve) illustrates the changes in supply that are likely to result from a variation in the price of oranges, and DD′ (the demand curve) the changes in demand. The curves intersect at P and the equilibrium price PX is, therefore, 1½d. This will be the price ruling in the market under the conditions assumed in Table I. For 1½d. is the only price which balances demand and supply. At any price higher than 1½d. there is always an excess of oranges which can find a market only when they are offered for sale below the current price. At any price lower than 1½d. there will

DIAGRAM II.—The Orange Market.




Number of Oranges.

always be a shortage of oranges and buyers will make higher offers in order to obtain supplies. A low or high price at one point in the market will communicate itself—on the assumptions that we are making—to all other points in the market. Thus if dealings begin at some price other than 1½d., they may be expected to oscillate for a time round 1½d. until it is clear that at this price a sufficient number of buyers will be forthcoming to take the whole supply of oranges off the market and that the requirements of buyers can be met in full. The price of 1½d. will then become the uniform price throughout the market.[1]

**Changes in Demand and Supply.**

A change in price will come about whenever there is a displacement of either of the curves. The demand curve, for example, may be moved

[1] This is not intended as a realistic description of how prices are fixed in a market like Covent Garden. The market spoken of above is not a place but a network of dealings (see above, p. 24).

to the right, to the position $D^1D$ in Diagram II, if the public develops a greater liking for oranges and is ready to buy a larger quantity than before at any given price. The price of oranges will be forced up steeply, falling back later to $P_1X_1$ as increased supplies come on the market in response to the rise in price. Similarly the supply curve $SS_1$ may be moved to the right if, for example, a new kind of orange tree makes it possible to grow oranges at less cost and so increase the supply offered for sale at any given price. A new and lower price will be established at which consumers make larger purchases than before.

We must carefully distinguish between an increase in demand or in supply represented by a movement of the whole curve and an increase represented by a movement *along* the curve. Only the first kind of increase can initiate a change in price; the second is a reaction to a change in price and cannot be the cause of it. A change in demand in response to a change in price is not an independent change in the state of demand but simply the reflection of some antecedent change in the state of supply. If there are more oranges to be had for a shilling, then people will buy more. But their increased purchases are no proof of a change in demand.

Changes in price, in short, are initiated by a change either in demand or in supply conditions. This is represented by a displacement of one of the curves, which, unless compensated by an off-setting displacement of the other (a very unlikely occurrence) must force the price to a new equilibrium level. Put differently, if demand and supply are tending to get out of line with one another, they can be kept in line by an appropriate change in price. In the last resort it will be price that preserves equilibrium.

**Difficulties in the Construction of Demand and Supply Curves.**

It is not always possible to draw demand and supply curves.[1] How, for example, can we draw a demand curve for motor-cars? What price are we to plot against the number of motor-cars being sold to the public? Should it be the price of Austins or Rolls-Royces, of 10 h.p. or 90 h.p. cars, of second-hand cars "in working order," or brand-new models fitted with the latest gadgets? Or if we try to calculate an average price for the whole group, how are we to average a fall in the price of some models against a rise in the price of others if the sales of each model do not respond equally to a given change in price? These practical difficulties make it impossible to represent the state of demand for motor-cars with complete accuracy in a demand curve. A demand curve is intended to represent the demand for something uniform—for *a* commodity; but motor-cars are a group of commodities rather than a commodity. Inside the group are cars which, from the point of view of producer or consumer, are as far removed from one another as a power-house is from a wireless set. A Bentley, for example, has more in common with an aeroplane

[1] The assumptions underlying the drawing of demand and supply curves will be examined in more detail below (pages 172 *et seq.*).

than with a Jowett. Other " commodities " are just as heterogeneous. It is rare for the products of a large number of firms to be exactly alike, and it is correspondingly rare, therefore, for a demand curve to be possible that represents demand conditions exactly.

A supply curve raises even greater difficulties. The curve is intended to show how much will be offered for sale at the market price. But what if there *is* no market price and each manufacturer can charge what he pleases ? Austin's do not first enquire after the price which their cars are fetching and then decide how many to sell. They try to *make* a market at a price which they fix themselves. They are willing to sell at that price far more cars than they are able to dispose of in practice ; and if sales begin to increase, they are just as likely to lower as to raise the price of their cars. A supply curve of Austin motor-cars, therefore, cannot possibly be drawn.[1] It is also impossible to draw a supply curve of motor-cars in general. If we cannot tell how many cars a particular firm will offer for sale at any given price, we cannot tell how many cars will be offered for sale by the whole industry. The fact is that the drawing of a supply curve presupposes keen competition between a large number of firms, each taking the market price for granted and with no power to control it. There must be no trace of monopoly in the market—not even such limited monopoly as arises when firms make goods differing ever so slightly from the products of their competitors.

The apparatus of demand and supply curves must always be used guardedly. It suggests to many people a permanency in the state of demand and supply which is not in keeping with the facts : both demand and supply are in a constant state of flux. It suggests, too, that it is price, and price only, that keeps demand and supply in line with one another—that price-competition is the only kind of competition. For many things, this is plainly not so. But, however one may question the legitimacy of drawing curves, they are an indispensable first step on the way to clear thinking about prices. The ladder may be rickety, and we may have to kick it out of the way later, but without it we will never reach the heights from which its shakiness becomes apparent.

**Elasticity.**

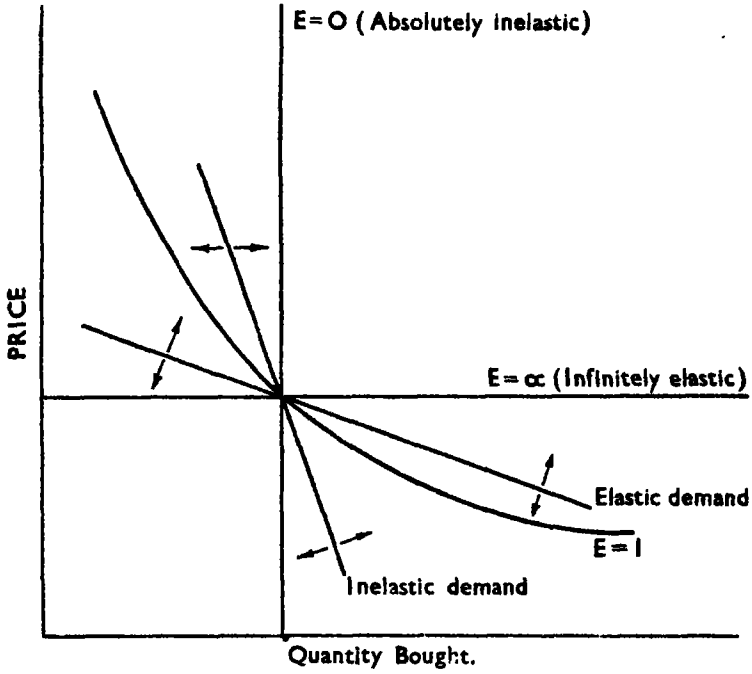
In order to understand how price balances supply and demand, we must study in more detail how each of these responds to changes in price. The rate of response is measured by what is called " elasticity." The elasticity of demand for a commodity is the rate at which the quantity bought changes as the price changes, other things remaining

[1] It is, of course, quite possible to draw a *cost curve* of Austin motor cars. But a cost curve and a supply curve are two quite different things. The first relates to the *firm ;* it is a curve of costs (i.e., of the expense to which a firm is put in order to produce a given output). The second relates to the *industry ;* it is a curve of *supply prices* (i.e., of prices, expectation of which will induce sellers to put a given quantity of their product on the market).

the same. The elasticity of supply of a commodity is the rate at which the quantity offered for sale changes as the price changes.

The term "elasticity" is used chiefly with reference to demand. Here it measures the ease with which people can put up with a small reduction in their consumption of a commodity, or alternatively the ease with which they can be induced, by a reduction in price, to consume a little more of it. If people buy the same amount of a commodity irrespective of the price, i.e. if they cannot do without *any* of the amount which they are buying, then demand is absolutely inelastic (elasticity is equal to zero). If people cease to buy the commodity

DIAGRAM III.—Elasticity of Demand.




altogether when it rises slightly in price, then demand is infinitely elastic (elasticity is equal to infinity). These are the outer limits. (See Diagram III.) Within these limits we can draw a dividing line between demands that change more rapidly, and demands that change less rapidly than price. Just on the dividing line, a given change in price will lead to an exactly proportionate change in the quantity bought; elasticity of demand is then equal to unity. On one side of the line, elasticity will be greater, and, on the other side, less than unity. When elasticity is greater than unity—that is, when a given change in price leads to a more than proportionate change in demand—we say that demand is *elastic*. When elasticity is less than unity—that is, when a given change in price leads to a less than proportionate change in demand—

we say that demand is *inelastic*. It will be observed that elasticity is a matter of degree and that, even when demand is inelastic, there are still some elements of elasticity.[1]

A second criterion (in the end, identical with the first)[2] is whether people spend more, or less on the commodity as it falls in price. If they spend more, demand is elastic; if less, demand is inelastic. In the borderline case, where elasticity is equal to unity, people will spend the same sum of money as prices fall, buying more of the article, but paying less per unit, so that the fall in price is exactly balanced by the increase in purchases.[3]

A rough estimate of elasticity can often be made by drawing what is called a "statistical" demand curve. In its simplest form, this consists of a number of points plotted from observed prices and sales in past years. Suppose, for example, that the potato crop and the price of potatoes in each of the last ten years is known. Then for each year we can plot one point and these points will lie roughly on a curve. From the shape of this curve, which will represent the response of sales of potatoes to variations in price, the elasticity of demand for potatoes can be calculated.

**Elasticity and Substitutes.**

The elasticity of demand for a commodity depends mainly on the range of available substitutes. The better the substitutes available, the more elastic will be the demand for the commodity. If motor-cars become dearer, the demand may not fall off very much, because people do not consider bicycles and buses very good substitutes for motor-cars. But if the price of Morris Tens goes up and there is a very similar model on sale (e.g. the Vauxhall Ten), demand will be transferred to the cheaper model and the sales of Morris Tens will fall heavily; that is, demand will be elastic. In the same way, the demand for tea will be less elastic than the demand for a particular

[1] The different possibilities can be illustrated diagrammatically.

[2] This can easily be verified by taking arithmetical examples. Students are sometimes puzzled by the fact that elasticity for a rise in price is different from elasticity for a fall in price. Suppose, for instance, that 9 units are bought at 8d. and 8 units at 9d. Then elasticity of demand is 8/9 for a fall in price and 9/8 for a rise; while if we adopt the second criterion based on expenditure, elasticity is unity. The difference between these values becomes negligible when we take two prices very close together. When the two prices differ infinitesimally, we get a single value for elasticity, measured at a price, not over a range of prices.

[3] It is instructive to turn back at this point to Table I and compare the elasticities of demand of different classes of orange-eaters. Working-class consumers have a consistently elastic demand. They spend more money on oranges at each reduction in price. Middle-class consumers spend more until the price is 1½d., then rather less at 1d., and more again at ½d. That is, there is a range of inelastic demand at prices between 1½d. and 1d., this range being wedged in between two larger ranges of elastic demand. Finally the well-to-do have an elastic demand at prices above 2d. and an inelastic demand below 2d.

brand of tea, because brands of tea are better substitutes for one another than cocoa, coffee, etc. are for tea. If there are absolutely perfect substitutes on sale (e.g. wheat of the same grade sold by different farmers), then the demand will be perfectly elastic; no one will be able to raise his price above the price paid to his competitors without losing his entire market.

The possibilities of substitution will be greater when there is a wide, range of uses to which the commodity can be put. Electricity, for instance, is used for heating and lighting, for the supply of power, and so on. Hence it can be substituted for coal, candles, gas, petrol, wireless batteries, etc., whenever it becomes cheaper. As electricity falls in price, people find it worth while to reduce their expenditure on a great many different things and transfer this expenditure to electricity. Thus the demand for electricity is fairly elastic. On the other hand, the demand for wheat is highly inelastic. Most people put wheat to a single use (i.e. they eat it), so that when wheat falls in price, they can only increase their purchases of wheat if they elect to eat more of it. This generally means a change of diet which is not very acceptable. So far, however, as wheat is used for other purposes (e.g. as a feeding-stuff) and so far as people are prepared to switch to wheat from potatoes, rye, oats, rice, and other substitutes whenever wheat is relatively cheap, the demand will retain some elements of elasticity.

Again, the demand for necessities is generally much less elastic than the demand for luxuries. The demand for potatoes, for instance, is far more inelastic than the demand for oranges. For necessities like salt or sewing thread, on which we spend a very small fraction of our income, demand will be particularly inelastic, since a rise of perhaps 50 per cent. in price may mean no more than a few pence on our annual outlay. Similarly, rich people have a less elastic demand than poor people. The poor man has constantly to be considering how he can spend his income to best advantage and is sensitive to changes in price; while the rich man, faced with a rise in price, may not bother to look for a substitute. The demand for pineapple will be far less elastic in Park Avenue than in Chinatown.

Elasticity arises in two different ways. Existing consumers will tend to buy more as prices fall, and new consumers will be tempted to make purchases. Sometimes, of course, only one source of elasticity is present. Practically everyone is already buying potatoes; new consumers cannot be found. On the other hand, it is rarely possible to persuade people to buy two copies of the same book; new consumers have to be found. Generally speaking, elasticity arises less through increased purchases by existing consumers than through the tapping of successive layers of demand from *potential* consumers as prices are brought down.

The conception of elasticity is of great practical importance. If the government is contemplating a reduction in the price of postage stamps or telephone calls: if the railways are thinking of reducing fares: if a manufacturer is trying to work out the best price for a new line in silk stockings—they must all form some estimate of the elasticity of

demand. Of course, they may not think in the precise terms in which elasticity has been formulated by economists. But there is no reason to suppose that their thinking is any the more accurate for neglecting (or avoiding) the economist's formulation.

**Elasticity of Supply.**

The elasticity of supply of a commodity measures the ease with which producers can meet a rise or fall in price by increasing or reducing supplies. Supply is said to be elastic or inelastic according as a change in price causes a more than proportionate or less than proportionate change in supply.

The elasticity of supply depends upon the range of alternatives open to the producer. If, for example, he is selling in several different markets his goods are likely to be in elastic supply to any one market; a fall in prices in that market will induce him to sell his goods elsewhere. Again, if he is producing several different goods, and can switch fairly easily from one to another, then each of his products will be in elastic supply. Or if the alternative of closing down his works and going out of business altogether will not involve him in heavy loss, then again supply is likely to be elastic. Finally, supply will be elastic if each producer's employees can readily obtain employment in other industries, if the materials which they use have alternative markets, and if their equipment is readily convertible to other uses. The more attractive these alternatives, the less will be the incentive to continue in the industry when prices fall, and the greater, therefore, will be the contraction in output and in supply.

With supply as with demand, elasticity arises in two ways: through a change in the output of existing producers and through a change in the number of producers. A full discussion of these changes must be postponed until later.[1]

**Fluctuations in Prices.**

Elasticity, whether of demand or of supply, makes for stability of prices. The price of a commodity like wheat, which is in inelastic demand, will have to fall drastically if a surplus has to be absorbed; whereas a comparatively slight reduction in price will be sufficient to secure the absorption of a surplus of oranges, which are in elastic demand. The more elastic the demand, the more stable the price. Similarly the price of commodities like rubber, which are in inelastic supply, will fluctuate more than the price of commodities like motor-cars, which are in elastic supply. The more elastic the supply, the more stable the price. Where demand and supply are very elastic—i.e., highly sensitive to changes in price—it will require no great change in price to keep them in step with one another. But where demand and supply are both very inelastic, and both extremely variable, the range of price fluctuations may be enormous. Potatoes, for example, may be 10s. a ton one year and £10 a ton the next, because large fluctuations in the crop encounter a highly inelastic demand.

[1] See Chapter 14.

# CHAPTER 13

# DEMAND

## § 1

PRICES are governed by supply and demand: that is, by cost on the one hand and utility on the other. The influence of cost on price is obvious enough, but not the influence of utility. What possible connection is there between the utility of water or of diamonds and their price? An answer to this question was put forward round about 1870 by three economists, each working independently of the others: Jevons in England, Walras in Switzerland, and Menger in Austria. The answer was based on a psychological generalisation —the Law of Diminishing Utility—on a new and very important conception—that of the margin—and finally on a logical distinction—the distinction between Total and Marginal Utility.

**The Law of Diminishing Utility.**

The value which we set on a commodity—its utility—depends upon how much we already have of it. The more we have, the less importance we attach to a further addition to our stock. If, for example, we have been smoking twenty cigarettes a day, we do not appreciate the offer of a free cigarette nearly so much as we should if our consumption had for some reason fallen to one cigarette a day. The utility of cigarettes is less when we are smoking twenty cigarettes a day that when we are forced to content ourselves with one.

**The Margin.**

Suppose now that cigarettes cost ½d. each. We are then faced with the problem, how many cigarettes to buy. We must weigh the satisfaction which we get from smoking cigarettes against the satisfaction which we might get by spending our money on other things. Each additional cigarette that we buy will have less and less utility, while each additional ½d. spent on cigarettes can be spared only at greater and greater sacrifice of other things. Ultimately our consumption of cigarettes must reach the point at which we are in doubt whether it is worth while to spend ½d. on another cigarette. We are then on the *margin of consumption* and the last cigarette which we think it just worth while to buy is the *marginal* cigarette. The margin is to be thought of as a boundary cutting off purchases which are worth while from purchases which are not. Some purchases will be well inside the boundary: we should never think of doing without at least one loaf of bread every week. Other purchases will be just as clearly beyond the boundary: we should never think it worth while to buy a hundred loaves of bread every week. On or near the boundary will be the purchases about which we hesitate—the fifth loaf of bread, the fourth packet of cigarettes, and so on. These are our marginal purchases.

The margin which bounds our consumption of any commodity is not fixed. A fall in the price of cigarettes will push the margin outwards towards purchases which were previously not worth while, but which become worth while at the lower level of prices. We may reject the twentieth cigarette as not worth ½d. and count the thirtieth cigarette well worth ¼d. The margin is drawn tightly or loosely according as things are dear or cheap; it is a boundary that shifts to and fro as prices move up or down.

**Marginal Utility and Total Utility.**

The total utility of a commodity is the utility of the whole amount which we are consuming at any given time; the marginal utility is the utility of the marginal unit—the unit which we think is least worth while to consume. The total utility will naturally increase with every addition to our consumption—unless we have already so much that further additions are a positive nuisance. The marginal utility will fall as our consumption increases since we shall be adding units which, in accordance with the law of diminishing utility, we appreciate less and less; as the margin of consumption is pushed outwards, utility at the margin, or marginal utility, falls. The total utility of a commodity measures the strength of our demand for the whole supply of it, the marginal utility measures the intensity with which we want a little more of it.

**The Paradox of Value Explained.**

Let us apply this distinction to water and diamonds.

The total utility of water is infinite: it would be impossible to give up using water altogether. But the marginal utility of water is negligible: we are put to no great inconvenience by having to do without a single pailful of it; and if we were given an *extra* pailful after using as much water as we required, it would clearly have no utility whatsoever. On the other hand, if everyone had to do without diamonds, there would be no great hardship, the total utility of diamonds is small in comparison with the total utility of water. The utility of a single diamond is, however, quite appreciable. Diamonds are so scarce that one more or less can make a considerable difference to our enjoyment of life. It is for this reason that diamonds are extremely valuable while water is free. The intensity with which people want one more gallon of water bears no comparison with the intensity with which they want one more diamond.

**Marginal Utility and Price.**

The idea of marginal utility gives us the clue to the connection between utility and price. We find, in fact, that the price of any commodity is governed, on the side of demand, by its utility at the margin of consumption, i.e. by its marginal utility. Suppose, for example, that we are trying to decide how many cigarettes to smoke per week. The larger the number of cigarettes that we smoke, the less is the marginal utility of cigarettes, and the less, therefore, is the

value that we set on an extra cigarette. But we cannot set one value on the last cigarette and a different value on the others. They are physically indistinguishable, and so of equal utility: and each costs exactly the same as the others. The price that we are prepared to pay for cigarettes, therefore, must be equal to the price of the last or *marginal* cigarette which we think it just worth while to buy. Now the price of the marginal cigarette is exactly balanced in our minds against its utility. The price of cigarettes in general must also be balanced, therefore, by the utility of the marginal cigarette. So we reach the important conclusion that the amount of any commodity which is purchased at a given price will tend to be such as to make the marginal utility of the commodity to each purchaser equal to its price. Demand operates so as to make marginal utility and price equal to one another.[1]

**Marginal Uses.**

Cigarettes, from which the marginal principle has so far been illustrated, have only one use. Most commodities, however, have several different uses which can often be arranged in order of importance. If there is only a limited supply of the commodity, this supply will be reserved for the more important uses; as the commodity becomes more and more abundant, the less important uses will be provided for one by one. With any given supply, or at any given price, there will be some uses to which it is just worth while to put the commodity: these will be marginal uses. Suppose, for example, that the amount of water is so reduced that we need to economise it and are willing to pay a price for more. For the sake of concreteness, suppose, that we are living in one of those villages in the North of India where the number of wells is small and the women draw water all day until at night the wells are dry. A shrinkage in the water supply will cause the margin of uses to which water is put to contract. First of all, people will be forced to give up using water for some comparatively unimportant purposes, e.g., irrigation. If more wells cease to give water, we may require to give up washing—first clothes and dishes, then faces and hands. Later we may have to ration drinking water. As the margin of uses to which water is put contracts, the utility of water in its marginal use mounts higher and higher. We are put to a greater and greater sacrifice by having to do without water first for one purpose, then for another. Conversely, the more water we have, the more trivial the uses to which we can afford to put it, and the less the price which we will be prepared to pay for more of it.

The utility of a commodity in that use which is just marginal tends, like marginal utility, to be equal to price. If the margin of consumption coincides with one particular use for the commodity, then utility

[1] This reasoning assumes that the market price is fixed and that consumers have no bargaining power. The price of cigarettes will cease to be equal to their marginal utility if the purchase of an additional cigarette affects the terms on which the remainder can be purchased.

in the marginal use is simply a special case of marginal utility. The price will be just sufficiently low to make this marginal use worth while, but sufficiently high to rule out uses of less importance to us. A fall in prices will cause the margin of uses to be pushed outwards; a rise in prices will make it press inwards again. Generally when a commodity becomes more plentiful, it is not only the margin of uses, but also the margin of consumption *in each use* that is pushed outwards. The margin is a boundary cutting each line of advance in consumption and not just a limiting point along one line. At all points on the boundary a marginal extension has a utility less than the price of the commodity, while a marginal contraction results in a sacrifice in utility greater than the price of the commodity. The margin is drawn round those units of the commodity which—whatever the use to which they are put—it is just worth while to buy.

**Marginal Purchasers.**

Just as there are marginal purchases, so there are marginal purchasers. A book, for example, will find some buyers at 15s., a larger number at 10s., and a still larger number at 5s. At any given price, there will be some people who find the book just within their means. These are the marginal purchasers—people who are on the margin of doubt whether they should buy the book or spend the money on something else.

**Consumer's Surplus or Rent.**

The fact that some purchasers are marginal while others are not, means that intra-marginal purchasers enjoy a surplus or rent. A consumer who would be willing to pay 15s. for a book, will be 10s. in pocket if he finds that the market price is only 5s In the same way, the fact that some purchasers are marginal while others are not, means that consumers enjoy a surplus on intra-marginal purchases. The marginal cigarette may be just worth ½d., but all the other cigarettes which we buy must, by the law of diminishing utility, be worth more. All cigarettes sell at the same price. If the marginal cigarette is just worth while, therefore, the other cigarettes must be more than worth while. They yield a surplus above their cost. Suppose, for example, that we have to choose between going without cigarettes altogether and paying a larger sum for our present consumption. Within limits, we should generally be willing to pay the larger sum. The difference between what we do pay and the maximum amount that we should be prepared to pay is our consumer's surplus on cigarettes.

**The Practical Importance of the Margin.**

There is an old saying:—"Take care of the pence and the pounds will look after themselves." This is simply a vivid application of the marginal principle. Proper attention to the margin will help consumers to get the most out of their limited income. Suppose, for example, that a housewife takes with her to the fruit market a basket and a purse, meaning to empty her purse and fill her basket with fruit.

How should she set about buying the best selection of fruit? Clearly, she must arrange her purchases so that she will not be better off if she spends more on one kind of fruit and less on another. This condition will be satisfied if the last pennyworth of oranges has the same utility as the last pennyworth of apples, bananas, and other fruits. Each addition to her stock of oranges reduces the marginal utility of oranges and raises the marginal utility of other fruits, since she has less money left over to spend on them. Her demand for oranges becomes less intense, while her demand for other fruit, becomes more intense. At the margin where these two intensities just balance, it is indifferent to her whether she spends her last penny on oranges or on any other fruit; there is no way in which she can gain by spending more on oranges and less on other fruits. Her money is spent to the best advantage (in the sense that her wants are most adequately satisfied) if she succeeds in equalising the marginal utility of money in each of its uses.

**The Theoretical Importance of the Margin.**

Analysis in terms of margins has the great merit that it fixes attention on the behaviour of buyers who are in doubt whether to buy a *little* more or a *little* less. Now it is the decisions of these persons that have most influence on the amount bought and so on prices. Those to whom "price is no consideration," or who buy without inquiring what things cost are not the "key" buyers. They affect prices in the sense that if they did not buy, total demand would be less, and prices, therefore, would be different; but they do not affect prices by boycotting what is dear and lining up where things are cheap. Their weight is thrown neither on the side of dearness nor of cheapness. It is to those who consider carefully what they can afford, who always ask the price, and constantly readjust their budget to stretch it to its furthest limits—to those, in other words, who, for whatever reasons, have an elastic demand—that we must look to keep down prices. Theirs are the marginal purchases which would cease to be made if prices rose. And it is at the margin that pressure on exchange values is exerted.

**The Margin, Elasticity, and Substitution.**

The idea of the margin is closely related to an earlier conception—that of elasticity. It is at the margin, and only at the margin, that there is elasticity. It is our marginal purchases that are discontinued, and marginal purchasers who cease to buy when prices rise. The more purchases and purchasers are marginal or near marginal, the greater will be the elasticity of demand. The two ideas, indeed, are held together as in a vice by a third idea—substitution. Elasticity, as we have seen, depends largely on the range of substitutes; and all substitution is essentially marginal. Substitution takes place at the margin of consumption; as prices move up or down we try to substitute a little of one commodity for a little of another. The more successful we are, the more elastic is demand and the

more readily does the margin shift in response to a change in price.

**The Law of Diminishing Utility Again.**

The ease with which the margin can be shifted depends upon the rate at which marginal utility is diminishing. So we have come round again to the law from which we started—the law of diminishing (marginal) utility. This is the fundamental law of demand and merits more attention, therefore, than we have given to it so far.

Stated formally the law is that, other things remaining the same, an increase in a person's rate of consumption of a commodity reduces the intensity with which further increments are demanded—that is, reduces the utility of the commodity at the margin. This is a law of almost universal application. It can be applied not only to things like bread and butter, railway journeys, men's hats, and so on, but also to the lectures of economists, the speeches of politicians, and even the number of suspects in detective stories. There are, however, some important limitations to the law. First, other things must remain the same. Suppose, for example, that a man is drinking beer. Then if each successive half-pint is to have less and less utility he must drink his beer, the same brand of beer, in the same public-house, with the same cronies, and the same landlord to judge his sobriety. He is not allowed to go wandering about enjoying a continual change of *milieu*. Secondly, we must not be trying to make a collection of something. If our beer drinker gets it into his head that he must complete a tour of the local public-houses, then the law of *increasing* utility will apply, for each successive drink will bring him nearer to his goal. Thirdly, our income, or at least our consumption of *other* commodities must remain constant. The marginal utility of beer depends upon how much wine we can afford. Similarly, the marginal utility of bread would increase if we could no longer afford to buy scones, tea-bread, and other wheaten foodstuffs; the marginal utility of a visit to the cinema would be much higher if we could not afford a wireless set, or fuel and light for our living-room. Finally, we must not develop a craving. A man who takes to drinking beer may find his thirst growing by what it feeds on, so that the more he drinks the more he wants to drink. Here again we have increasing, instead of diminishing, utility; although since the man is no longer the same, so to speak, after drinking as he was before, it is a little difficult to decide for *whom* beer has increasing utility—the novice who took to drink or the drunkard with the craving for it.

The development of a craving is a special case of what quite commonly suspends the operation of the law of diminishing utility—a change of tastes. We may begin, for example, by detesting symphony concerts; the marginal utility of a symphony concert is then negative and seems likely to become increasingly negative the more symphony concerts we are forced to listen to. In time, however, by dint of attending these concerts, we may find that we

have developed quite a liking for this kind of music and look forward to hearing more of it. An increased consumption of symphonies may raise our appreciation of them and hence also their marginal utility.

**Other Laws of Demand.**

The psychology of desire is too complicated to be embraced within a single law. Of all aspects of human desire, satiability and love of variety are most relevant to economics; and both these aspects are expressed in the law of diminishing utility. But the law takes no account of other aspects of desire—the part played by impulse, or example. Again, it concerns only the form and not the content of our wants: in a full study of demand we should require to analyse the things that shape each of our wants and purchases—our love of power, our desire for a rôle in life, our fear of boredom or of social disapproval, the necessity to keep up with the Joneses, and so on. But an attempt to consider these motives and to reduce them to general laws of demand would take us much too far afield.

§ 2.

**An Alternative Version.**

The theory of marginal utility, even when it is stated in a simplified form, is difficult for students to understand. The word "utility" leads to constant confusion. First of all, "utility" is used in economics to mean "desiredness" or "strength of demand." But it is associated in people's minds with "usefulness" and "desirability." The result is that "utility" can be used to muddle up economic and ethical considerations. This was repeatedly done by Jevons, Edgeworth and others, and is still being done by professional economists. Students are to be excused, therefore, if they fall into a confusion which besets even the most wary. Secondly, the theory of marginal utility was worked out at a time when most economists were utilitarians[1] (hence the muddle of economics with ethics) and believed that utility was something quantitative. They spoke of a quantity of utility as they might have spoken of a pound of cheese. But it would be as reasonable to speak of a quantity of temperature. The fact is that we can only measure the intensity of one demand in terms of the intensity of another, and that, behind the screen of "utility," is exactly what the theory does—or should do.

If we remove the screen, what kind of theory are we left with? In its essentials the new theory is very much like the old. But it lays emphasis on the word "preference" instead of on the word "utility." Where we talked of "marginal utility" we now talk of "marginal preference"; and where we drew curves of marginal utility we now draw "indifference curves."

[1] i.e., they measured the goodness of an action by its utility.

**The Margin : Preferences.**

When we value an article we do so only in relation to the alternatives open to us—in relation, for example, to the available substitutes. Our values are the expression of a preference. Now these preferences may be absolute. We may detest cocoa and adore tea. If asked to choose between the two we may always plump for tea. This rules out one possible alternative to tea and narrows the field of choice. But there are still plenty of other alternatives where, just because our means are limited in comparison with our wants, we *cannot* have absolute preferences. We are forced to balance tea against these alternatives and to *modify* our preferences if the alternatives become less attractive or more attractive because of a rise or a fall in their price. Suppose that we group these alternatives together and call them, for short, "coffee." It may happen that if tea costs twice as much as coffee, we give up drinking tea, while if coffee costs twice as much as tea, we give up drinking coffee. Within this range of prices we will probably drink both tea and coffee, but will vary the proportion in which we drink them so as to take advantage of changes in the relative price of tea and coffee. We may find it possible to draw up a schedule showing the various combinations in which, if our income remains fixed, we will drink tea and coffee at different price-ratios. The schedule may be something like this :—

| Cups of tea | Cups of coffee | Price of tea / Price of coffee |
|---|---|---|
| 0 | 50 | 3·0 |
| 3 | 41 | 2·5 |
| 5 | 37 | 2·0 |
| 9 | 30 | 1·5 |
| 14 | 23 | 1·25 |
| 21 | 15 | 1·0 |
| 35 | 3 | 0·75 |
| 40 | 0 | 0·5 |

This schedule illustrates three points. First, that we give up coffee for tea, the dearer coffee becomes in relation to tea. A change in price leads us to substitute the one for the other. Second, that the number of cups of coffee that we are willing to sacrifice for the sake of *one more* cup of tea decreases the more cups of tea we drink. Rather than do without our first three cups of tea, for instance, we are prepared to give up 50 — 41 = 9 cups of coffee. Each cup of tea, to begin with, is worth three cups of coffee to us. But when we are already drinking 14 cups of tea, we are prepared to sacrifice no more than eight cups of coffee for seven more cups of tea—that is, a cup of coffee and a cup of tea are of much the same value to us.

Thus we find it progressively harder to substitute a little more tea for a little less coffee, the more tea and the less coffee we are drinking already. Our preference for tea over coffee, *at the margin* where we have to balance a little more of one against a little less of the other, falls steadily. Given the amounts of tea and coffee that we are drinking, this "marginal preference" can be measured by the ratio in which we are prepared to sacrifice a small quantity of coffee for the sake of a little more tea. In the two examples given above, our marginal preference is approximately 3 and 1. Clearly our marginal preference is not the same thing as our general or total preference; just as marginal utility is not the same thing as total utility. We may prefer tea to coffee in the sense that we should prefer to do without coffee altogether rather than do without tea altogether. But we may simultaneously value an occasional cup of coffee twice as highly as our usual cup of tea. If a cup of coffee were to cost twice as much as a cup of tea, our general preference for tea would not necessarily prevent us from buying coffee now and again.

A third point illustrated by the schedule is that the relative price of tea and coffee and our marginal preference for one in terms of the other tend to equality with one another. As relative prices fall from 3·0 to 0·5, marginal preference falls from 3·0 to 0·6. The coincidence is not exact, but this is because the schedule is not sufficiently detailed, so that marginal preference has to be measured by taking comparatively large variations in our consumption of tea and coffee, instead of the small (really infinitesimal) variations assumed in pure theory. Why relative prices and marginal preference should tend to equality is not difficult to understand. Marginal preference measures the terms on which we are prepared to sacrifice a little of one commodity in order to obtain a little more of another. Relative prices measure the terms on which we *can obtain* (by exchange) a little more of one commodity by parting with a little of another. If the two are not equal, therefore—if our marginal preferences are out of line with relative prices—we must be refusing to strike an advantageous bargain, or overlooking an opportunity of gain. If, for example, a cigar can be had for the price of 10 cigarettes, and we count the satisfaction of one more cigar equal to the sacrifice of 12 cigarettes (i.e., if relative prices are in the ratio of 10 : 1, and our marginal preference is 12 : 1), then the purchase of another cigar in place of 10 cigarettes leaves us two cigarettes to the good. We will switch, therefore, from cigarettes to cigars until our marginal preference is brought into line with their relative price—until, in fact, we are indifferent whether or not we use the price of a cigar to buy cigarettes.

We modify our preferences, then, so as to bring them into line with prices. But this is only part of the truth. Prices are themselves pulled into line with our marginal preferences. Demand, as we saw above, not only responds to changes in price but helps to determine price. An increased supply of mutton, for example, will find a

market only at a price which is low enough to encourage the substitution of mutton for other commodities (e.g., beef)—at a price, in other words, which is in keeping with people's preferences for mutton as compared with beef (and other alternatives) at the margin of choice. If marginal preferences are inelastic, then the fall in mutton prices will be heavy. But if marginal preferences are elastic (i.e., if a small change in relative prices is sufficient to cause a large switch-over in consumption) then the extra supply of mutton on the market will be absorbed without any drastic cut in prices.

Now marginal preferences will obviously be elastic if there are plenty of satisfactory substitutes. We will have no strong preference for a commodity if a wide range of alternatives is available, and we will be able to extend or contract our consumption of it whenever one of these alternatives becomes slightly dearer or slightly cheaper. The elasticity of our preferences, therefore, just like the elasticity of demand, as defined above, depends upon the range of available substitutes. This is not surprising, for elasticity of demand is just an example of elasticity of preference—the example *par excellence*. Suppose, for example, that we are measuring the elasticity of demand for butter. What we are really doing is to weigh up our preference for a little more butter against the sacrifice of a little more money and to find out how elastic this preference is as the cost of butter in terms of money changes. Money sums up all the alternatives to butter and not just a few, like margarine or jam. It stands for all the substitutes that we might buy if butter were to become dearer. Just as money is the common denominator of exchange, so elasticity of demand is the common denominator of elasticities of preference. We can convert all exchange values into money values; and in the same way we can convert our preference for one commodity as compared with another into our preference for each as compared with money.

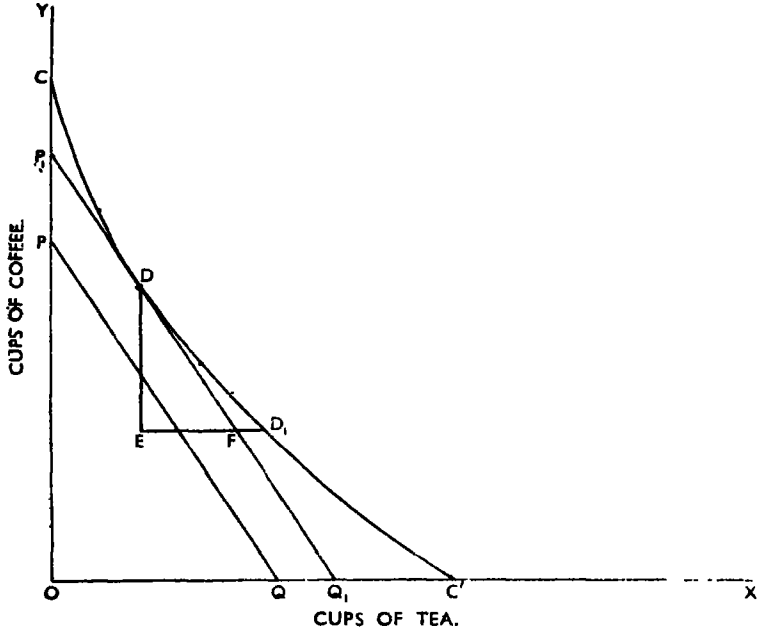
If one price is high and another low, therefore, the conclusion to be drawn is, not that we prefer the first commodity to the second, but that we prefer *a little more* of the first to *a little more* of the second. The value that we set on commodities is governed by our preferences at the margin of choice, i.e., by our marginal preferences. If these preferences are inelastic, so that only a small proportion of our purchases are marginal, then comparatively small changes in supply will cause very large changes in prices. If, however, our preferences are elastic, so that a large proportion of our purchases are marginal, prices will be comparatively stable. The stability of prices depends upon the ease with which we can vary our demand to meet a change in the pressure of supply. It is at the margin that this change in pressure must be studied. The pressure of demand on prices depends upon our marginal preferences; the pressure of supply, as we shall see in the next chapter, depends upon marginal costs.

**Note on Indifference Curves.**

The schedule given above can be plotted on a graph (see Diagram IV)

cups of tea being measured along the X axis and cups of coffee along the Y axis. The resulting curve is known as an indifference curve, since any point on it represents a combination of cups of tea and coffee on which we set equal value. The actual combination that we select depends upon the relative price of tea and coffee. This relative price can be represented by the slope of a straight line falling from left to right (PQ in the diagram). Any point on this line will represent a combination of tea and coffee of constant value at current prices. (OP cups of coffee, for example, are necessarily equal in value to OQ cups of tea, since $\frac{OP}{OQ}$ measures the relative value of tea and

DIAGRAM IV.—Indifference Curve: Tea—Coffee.




coffee). Marginal preference can be represented by the slope of the tangent at any point to the indifference curve $CC^1$, for this slope represents the rate at which we are prepared to sacrifice a little coffee for the sake of a little more tea. The combination of tea and coffee selected will tend to be one that equates our marginal preference and relative prices—that is, it is represented by the point (D in the diagram) at which a line parallel to PQ touches $CC^1$.

The combination selected must lie on $CC^1$; it cannot lie to the right because our income is too limited, nor to the left because we should then be spending our income carelessly. It is clear, too, that the only point on the curve that we are free to choose, at prices represented

by the slope of PQ, is D. The line PQ represents combinations of tea and coffee of equal value at current prices, and any point lying above PQ, therefore, involves an increase in outlay. Take, for example, the point $D_1$. We should be willing to give up ED cups of coffee in order to increase our consumption of tea by $ED_1$ cups. But at current prices we can increase our consumption of tea by EF cups only. The remainder, $FD_1$, represents additional expenditure that is necessary before we can move to $D_1$; and once there, we are no better off than at D. There is thus a pure gain of $FD_1$ in remaining at D.

If prices change, then the slope of PQ changes also, and we move to a new point on the curve. To each price-ratio there corresponds a point that represents the optimum combination of tea and coffee at that ratio. Each point represents an outlay which, at the appropriate price-ratio, is of constant value. The tangent at each point measures our marginal preference at the level of consumption which the point represents. As the tangent changes slope, there is a change in the proportion in which we combine tea and coffee. If the second change is rapid in comparison with the first, our elasticity of preference or (as it is generally called) of substitution is high. But if the change is comparatively gradual, if it requires a great change in relative prices to make an appreciable change in the proportion in which we consume tea and coffee, then our elasticity of preference, or of substitution, is low. Elasticity will be infinite if the indifference curve is a straight line. For example, two commodities, which are exactly alike except in size, will presumably sell at fixed prices in terms of one another, and it will not matter how they are combined with one another. The slightest rise in the price of one will cause everyone to buy the other. Elasticity will be zero when one commodity cannot be substituted for another. The indifference curve will then shrink to a single point, and marginal preference will be indeterminate.

# CHAPTER 14

# SUPPLY

## (1) COMPETITION AND MONOPOLY

IN the last chapter we saw how the price of a commodity is governed on the side of demand, by the value that consumers set on marginal units of it. We have now to see that, on the side of supply, the price of a commodity is governed by the cost of producing marginal units of it. Just as we tried to go behind the demand curve in order to study the preferences of consumers, so we must now try to go behind the supply curve to the costs of producers. We shall find that the link between cost of production and price is much more complicated

than the link between utility, or preference, and price. Cost of production is not nearly so simple a conception as it seems.

Cost controls supply in two ways. It controls the volume of output which each firm finds it profitable to produce, and it controls the number of firms that can carry on at a profit. If the cost of producing a commodity rises, therefore, the supply will contract for a double reason—first, because each firm will cease to manufacture units of output that no longer pay their way; and, second, because some firms will find it necessary, or advantageous, to abandon production of the commodity altogether. The rise in costs will press on supply at the margin—on units of output which it was just worth while for each firm to produce, and on firms which found it just worth while to carry on production of the commodity. Other intra-marginal units of output will continue to be supplied, and other (intra-marginal) firms will continue in business at the higher level of costs.

Thus it is marginal, rather than average, cost that controls supply. Marginal cost is the net cost of a marginal addition to output. This addition may be made either within the individual firm or through the entry of a new firm into the market. Marginal cost, therefore, can be used in two quite distinct senses. It can be used to mean the cost to any one firm of increasing its output by a single unit. Or it can mean average cost of production in that particular firm which just finds it worth while to continue production. Marginal cost in the first sense may be quite different from marginal cost in the second sense.

It might be supposed that just as price tends to be equal to marginal utility, so price would tend to be equal to marginal cost (in one or other of the above senses) This supposition would be correct if it were possible to make assumptions about supply parallel to the assumptions that were made implicitly in our study of demand. We assumed a state of competition between consumers in which each bought only a small quantity of the total supply, had regard solely to his own interest, and had no power to influence prices. Consumers, that is, were supposed to take prices for granted, and to be able to buy unlimited quantities of any commodity at the ruling market price. These assumptions, which are not altogether unrealistic when applied to demand, are generally quite inappropriate when applied to supply. One producer very often sells a substantial proportion of the total supply; his competition with other firms may be limited by understandings or agreements; and his power over prices may be far from negligible. Producers, that is, are rarely able to take prices for granted and sell as much as they please at the ruling price. They have to regulate their output so as to keep from forcing down prices and spoiling their market. Similarly, a new firm may be unable to find a market without incurring special selling costs (e.g., by advertising). It has to build up a connection and "make a market" for itself, instead of simply offering its goods for sale at current prices.*

In the special circumstances in which competition between producers is similar to the competition which we supposed to exist between consumers—when competition is " perfect "—price and marginal cost (in both senses) will be equal. In all other circumstances price will be above marginal cost. The importance of a clear understanding of these two propositions and all that they imply cannot be exaggerated.

**Price and Marginal Cost : (1) The Individual Firm.**

It is convenient to begin by assuming, for the sake of simplicity, that the firm is making a single commodity only, and that it does not resort to advertising or to any kind of sales pressure. These assumptions (which are obviously unrealistic) rule out a number of complications for which provision can be made later.

The firm will be likely to aim at making as large a profit as possible. This aim will be realised when two quantities—marginal cost and marginal revenue—are equal. Marginal cost, as we have seen, is the net cost of a marginal addition to output; marginal revenue is the net revenue obtained from the sale of a marginal addition to output.[1] Or, to speak more precisely, marginal cost represents the increment in aggregate costs when output is increased by one unit—or the expense that could be avoided by reducing output by one unit. Marginal revenue represents the increment in aggregate receipts when one more unit is sold—or the reduction in aggregate receipts when one unit less is sold.[2] Now if the firm is producing that output at which its profits are maximised, an increase or decrease in output by one unit must leave profits lower than before. In other words a marginal addition to output will increase aggregate costs by more than it increases aggregate revenue, while a marginal reduction in output will save less in cost than it sacrifices in revenue. Marginal cost cannot be above marginal revenue unless at least one unit of output is not paying its way; nor can it be below marginal revenue or there will be an extra profit to be made by making and selling at least one

[1] Students must on no account confuse revenue with profit. The revenue of a firm is its receipts from the sale of goods; the profit of the firm is the surplus of these receipts over the cost of producing the goods.

[2] It is obvious that marginal cost and marginal revenue may be greater or less in value when output is being expanded than when output is being contracted. As a rule the difference will be insignificant. But if, for example, a change has to be made in the organisation of the firm, or in the type of plant in use, before output can be expanded, there will be a discontinuous change in costs and revenue at that level of output. Marginal cost and marginal revenue for an expanding output may be high, while for a declining output they are comparatively low. An additional complication is introduced, as we shall see in discussing overhead costs, by the time factor. The expense that is avoided by reducing output by one unit is not the same in the short period as it is in the long run. No overhead costs can be avoided in the short period, but they *can* be avoided in the long run.

more unit of output. Marginal cost and marginal revenue must, therefore, be equal when profits are a maximum.[1]

The meaning and implications of this statement are illustrated in Table 2. Suppose that the Scribblo Propelling Pencil Company has

TABLE 2.[2]

*Scribblo Propelling Pencil Company, Ltd.*

| Number of pencils (millions) | Average cost (pence) | Total cost (million pence) | Marginal cost (pence) | Average revenue (million pence) | Total revenue (million pence) | Marginal revenue (pence) | Profit (million pence) |
|---|---|---|---|---|---|---|---|
| 1 .. | 28 | 28 | 28 | 36 | 36 | 36 | 8 |
| 2 .. | 20 | 40 | 12 | 30 | 60 | 24 | 20 |
| 3 .. | 17 | 51 | 11 | 25 | 75 | 15 | 24 |
| 4 .. | 15 | 60 | 9 | 21 | 84 | 9 | 24 |
| 5 .. | 14 | 70 | 10 | 17 | 85 | 1 | 15 |
| 6 .. | 15 | 90 | 20 | 14 | 84 | - - 1 | - 6 |
| 7 .. | 18 | 126 | 36 | 11 | 77 | — 7 | — - 49 |
| 8 .. | 22 | 176 | 50 | 8 | 64 | - - 13 | — 112 |

before it estimates of costs and sales similar to those in the table. From these estimates the company can calculate marginal cost and marginal revenue at each level of output. Thus, when three million pencils are being produced, marginal cost is (roughly)[3] 11d., since the last million pencils has increased total costs by 11 million pence. Marginal revenue at the same level of output is (roughly)[1] 15 pence, since the last million pencils sold has increased total revenue by

[1] The proposition that firms will earn their maximum profits by equating marginal cost and marginal revenue does not imply that business men have the perspicacity to do so. Few business men have heard either of marginal cost or of marginal revenue : and only a proportion of them have the costing machinery required for the estimation of average, still fewer for the estimation of marginal, cost. Even if costing were in universal use it would in many industries be extremely difficult to apply it to the estimation of marginal cost, while marginal revenue must almost always be a matter of guesswork. Nevertheless business men do at times work—not always *explicitly*—in margins and those who do not will tend to be weeded out because they habitually make lower profits than they might.

[2] The data in columns 1, 2 and 5 is estimated by the firm : from this data the figures in the remaining columns are calculated.

[3] A million pencils is much too large a variation in output to allow us to estimate marginal cost and marginal revenue accurately. We require to take two outputs that are very close to one another—ideally, two outputs that differ only infinitesimally.

15 million pence. Profits reach a maximum when output is 4 million pencils, and at this output marginal cost and marginal revenue are equal. It is true that profits are just as great when only 3 million pencils are being produced; since the final unit of output just pays its way, profits are not affected by its production. But the final unit in the table is very large (a million pencils), whereas with fuller knowledge of costs and sales it can be made negligibly small. If the final unit is infinitesimally small, as we assume in theory, then it is of no importance whether we say that output is such that marginal cost and marginal revenue are equal, or whether we say that output will be just one unit less than this.

Marginal cost and marginal revenue are rather elusive conceptions until they are pinned down and joined to the more familiar ideas of average cost and average revenue (or price). The typical relationship between these quantities is illustrated in Table 2.

**Marginal and Average Cost.**

At first the firm enjoys the benefit of economies of large-scale production, and costs fall as output increases. Marginal cost is below, and falls faster than, average cost. The reason for this is that each unit added to output yields economies in the production of all the preceding units. These economies must be deducted from the direct outlay on the marginal unit of product in order to arrive at marginal cost; but, in calculating average cost, the economies have to be averaged over the whole of output, and the fall in average cost as output expands is, therefore, less rapid. At a later stage, costs begin to rise, either because the firm is too large for efficient management, or because of increasing risk, or because output presses on the supply of some scarce factor of production, so that either a higher price has to be offered in order to attract more of it, or units of inferior quality or efficiency have to be used.[1] First, marginal cost begins to rise; but so long as it remains below average cost (i.e., so long as each additional unit costs less than the average) average cost continues to fall. Later, when marginal cost exceeds average cost (i.e., when each additional unit costs more than the average) average cost rises too. When marginal cost is exactly equal to average cost, average cost will be neither rising nor falling but will be at the turning point where it touches bottom.[2]

**Marginal Revenue and Price.**

Marginal revenue in Table 2 is consistently lower than price. Any attempt to increase the sale of pencils involves a reduction in price, and this reduction applies to all pencils, not just to those which are

[1] For instance, if more workers are needed, it may be necessary to offer higher wages so as to attract workers from other districts or other trades; alternatively, workers of less than average skill may have to be taken on at the standard rate of wages. For a fuller discussion, see Chapter 20.

[2] Thus if we draw curves of marginal and average cost, the marginal cost curve will intersect the average cost curve at its minimum point.

bought because of the cut in price. Thus the pencils previously disposed of at the higher price bring in less revenue than before, and this loss of revenue has to be deducted from the gain in revenue through the sale of additional pencils before we arrive at the net increase in total revenue. Now the net increase in total revenue is simply marginal revenue, while the gain from the sale of more pencils is equal to the price of pencils. Marginal revenue, therefore, must fall short of price by an amount equal to the concession that has to be made to existing buyers in order to extend sales. This concession is great if demand is inelastic and small if demand is elastic.[1]

If, for example, demand is perfectly elastic, then increased sales can be made without any reduction in price. No concession to existing buyers need be made and marginal revenue and price are equal. On the other hand, if demand is inelastic, marginal revenue must be negative, since consumers will spend less money on the commodity (although they may buy a larger quantity of it) as it falls in price. The concession made to existing buyers outweighs the additional revenue coming from increased sales.

The ratio of marginal revenue to price, then, is governed by the elasticity of demand for the product of the firm.[2] Elasticity in its turn, as we have seen, is governed by the range of available substitutes. But these substitutes are simply competing products. So we arrive at the conclusion that the ratio of marginal revenue to price measures

---

[1] If we assume that X pencils are sold at a price P, and that $(X+\Delta X)$ pencils are sold at a price $(P-\Delta P)$—where $\Delta X$ and $\Delta P$ stand for small changes in X and P—then marginal revenue is equal to :—

$$\text{Increase in Revenue} = (P-\Delta P)(X+\Delta X)-P.X$$

$$\text{Increase in Quantity Sold} = \Delta X$$

$$= P - \frac{X.\Delta P}{\Delta X} - \Delta P$$

The last item disappears when $\Delta X$ and $\Delta P$ are infinitesimal. Thus price exceeds marginal revenue by an amount equal to $\frac{X.\Delta P}{\Delta X}$

$$\text{But elasticity of demand} = \frac{\frac{\Delta X}{X}}{\frac{\Delta P}{P}} = \frac{P}{X.\frac{\Delta P}{\Delta X}}$$

Hence the excess of price over marginal revenue is inversely proportionate to elasticity of demand.

[1] Notice that we are talking about the demand for the product of a *firm*, not, as in Chapter 12, about the demand for a *commodity* (i.e., the product of an industry or group of firms competing with one another). For example, we are discussing the demand for Austin motor cars rather than for motor cars in general.

the severity of the competition which the firm has to face. By implication, since marginal revenue and marginal cost tend to equal one another, the ratio between marginal cost and price is an index of the severity of competition. The more nearly competition approaches "perfection," the closer does price come to (marginal) cost. The greater the monopoly power of a firm, the greater, in general, is the discrepancy between price and (marginal) cost. These conclusions require elaboration.

**Monopoly and Competition.**

The average man draws a sharp distinction between monopoly and competition. Under monopoly, he thinks, someone has complete, or at least substantial, control over the whole supply of a commodity; whereas under competition the commodity is supplied by several independent firms. This distinction is satisfactory only if we can draw a hard and fast line between one commodity and another. A monopolist, we say, is selling something different from other people. But how different? Must the difference be as great as that between milk and tea, or can it be as small as that between milk from one farm and milk from another, or between one brand of tea and another? Are we to call Bovril and Oxo and Camp Coffee commodities and their manufacturers monopolists? If a single firm supplies us with 8 h.p. motor-cars while other firms confine themselves to higher-powered cars, is the first firm a monopolist? The more we reflect, the more we are driven to conclude that practically every firm is in some sense a monopolist. Everywhere in business, firms are selling a product which is differentiated in style, brand, quality, package, size, etc., from the products of other firms. Each firm constantly lays stress on the uniqueness of *its* product, insisting that it is not only better but different from other products. Even where no obvious physical difference exists, firms try by advertisement to create the conviction that there is a real difference. They try also by the offer of special services to make it more convenient for customers to deal with them rather than with competitors. In a wide variety of ways, firms seek to insulate t m competition by creating a special market in which re preferred to others. They are able to bind custome s in such a way that they will neither lose all their cust rge a price appreciably higher than elsewhere, nor gain number of new customers by charging a price appreci eir discretion in fixing prices is generally limited, and be in broad correspondence with prices in other parts But the fact that firms have any power at all over pr at they have acquired powers of monopoly within their

The fact is that we nev oly undiluted by competition, and very rarely find co luted by monopoly. In most lines of business there mpetition and monopoly in which one or the other ate. The difference between monopoly and competiti ree, not of kind.

**Elements of Competition under Monopoly.**

Competition is fundamentally the offer of a substitute; monopoly is fundamentally the absence of substitutes. But there can never be an entire absence of substitutes. So long as our purchasing power is limited, everything on which we might be induced to spend our money is in a sense a substitute for everything else that we buy. If any one commodity becomes too dear we can always give up using it and substitute some other commodity. We may continue to satisfy the same want if there are close substitutes (e.g. margarine for butter), or we may be driven to satisfy some quite different want so that the substitution becomes very roundabout. What we obviously cannot do is to go on buying a constant amount of any one commodity, however high its price.

Even if competition within an industry is suppressed, therefore, the monopolised product has still to face the competition of other industries. Indeed, such competition as survives nowadays is more and more between industries and less and less between firms within an industry. As monopoly grows, prices are arrived at through bargains between organised industries, and the common interests of the firms in each industry come into open conflict with the interests of the allied firms in other industries. A particularly striking example of the change in the arena of competition from the firm to the industry is provided by those joint advertising campaigns in which one industry appeals to us to Eat More Fruit, a second to Eat More Bread, a third to Eat More Fish, a fourth to Eat More Margarine, and a fifth to Eat More Butter—invitations which are obviously competitive, since the more heartily we fall in with one the less our ability (in appetite and in purse) to fall in with the others.

The competition with which a monopolist is faced comes not only from other products already on the market but also from *potential* substitutes that might be made available to consumers if prices were to rise. High prices may drive consumers to devise new substitutes for the monopolised goods or tempt new firms to enter into competition with the monopolist. Such potential competition will tend to keep prices down, since the monopolist will have no wish to invite it and will have no confidence in his ability to meet it if he has lost the good-will of his customers by overcharging.

The power of monopolists [redacted] erefore, is limited by the competition of substitutes, ac [redacted] tial—in other words, by the elasticity of demand for [redacted] st's product. The more elastic the demand, the more [redacted] rice and (marginal) cost approximate to one another. [redacted] c the demand, the greater will be the power of the mono [redacted] more excessive the price which he is able to charge. S [redacted] o the conclusions reached above. Every producer is mo [redacted] a monopolist; and every producer will maximise his pr [redacted] ng marginal revenue and marginal cost. The more of [redacted] he is, the less elastic will be the demand for his produ [redacted] er, therefore, will be the

ratio of price to marginal revenue; and the greater, finally, will be the ratio of price to marginal cost.

**Perfect Competition.**

Perfect monopoly—unlimited control over the market price of a commodity—does not exist; perfect competition—competition purged of every element of monopoly—is rare, but it does exist. For competition in a commodity to be perfect, each individual buyer and seller of the commodity must regard the market price as entirely outwith his control. Buyers will have no incentive to restrict their purchases for the sake of driving down the price; sellers will have no incentive to restrict output or withhold supplies for the sake of maintaining or raising the price. Each buyer or seller will form his own judgment of the probable course of prices and will treat the price which he anticipates as unaffected by his individual purchases or sales.

Perfect competition on the side of buyers is common enough. The average housewife, competing with millions of other housewives, could not conceivably, by the impact of her purchases, budge the price of staple commodities. On the side of supply, examples of perfect competition are much harder to find. Wheat is one: the wheat farmer in the Middle West is hardly likely to hesitate about growing a few bushels more or less for fear of upsetting prices in the Chicago wheatpit. Gold is another example: the goldminer is unlikely to be afraid of spoiling his market by despatching an extra ounce or two of gold to London. Even in wheat and gold, however, competition is not altogether perfect. The price of wheat is greatly influenced by the operations of wheat pools, storage schemes, and so on, all of which introduce elements of monopoly into the market. The price of gold depends mainly on the willingness of a single buyer, the Federal Reserve Board to go on paying $35 an ounce to have gold dug up, and several million dollars to have it buried again. Thus it is only from the point of view of the individual farm or mine that competition in wheat and gold is perfect.

When there is perfect competition, price ceases to be above marginal cost. Each firm can sell freely as much as it chooses without fear of spoiling its market; it can treat the demand for its product as perfectly elastic. If the firm tries to increase its sales, the additional supply put on the market is negligible in comparison with the total supply, and the fall in price, therefore, is so small that it does not enter into the firm's calculations. Price and marginal revenue are then necesarily equal, since each additional unit sold adds its full price to the total revenue of the firm; and if the firm is successful in equating marginal revenue and marginal cost, so as to maximise its profits, price and marginal cost must also be equal.

Perfect competition can be regarded, therefore, as the limiting case of imperfect or monopolistic competition. The more nearly competitive conditions approach "perfection," the closer to unity will be the ratio of price to marginal cost; the greater the deviation towards

monopoly the further will price rise above marginal cost. How do these deviations arise?

**Elements of Monopoly under Competition.**

**1. Market Imperfection.**

We have seen that firms seek to shelter themselves from competition by creating a special market in which they enjoy the goodwill of their customers. The existence of goodwill, however, is incompatible with perfect competition. For the first requirement of perfect competition in any commodity is that the market in that commodity should be perfect: a perfect market being one in which buyers have no preferences as between the different units of the commodity offered for sale, sellers are quite indifferent to whom they sell, and both buyers and sellers have full knowledge of prices in other parts of the market. If markets were perfect, each firm would have to face the competition of products which were not only physically interchangeable with its own but which were considered by consumers to be perfect substitutes in every way. No firm would be able to charge more than the market price without losing the whole of its custom to its competitors, while no firm could charge less than the market price, since a price-cut by any one firm would force all the others to follow suit. There not only would be, but *could* be only one price throughout the market.

The markets of the real world are never perfect. First of all, there are imperfections due to ignorance. Buyers and sellers may be ill-informed about the terms on which dealings are proceeding elsewhere. Consumers, for example, generally know the price charged in a limited range of shops only, and buy from their regular suppliers without troubling to obtain other quotations. In the same way, they may go on buying a familiar brand in ignorance of cheaper substitutes on the market. Secondly, there are preferences due to the different suitability of competing goods for the purposes of different buyers. Each seller will try to cater for the special requirements of a given group of buyers, offering them goods of a quality, style, colour, etc., nearer to their taste than the goods offered by his competitors. Or he will try to carve out a special market by providing additional services—regular delivery, long credit, free insurance, gift coupons, etc. Some of his customers—those who are farthest away, or whose preferences are least satisfactorily met—will remain, so to speak, on the fringe. They will be *marginal*, in the sense that they can be most readily detached by competitors. Other customers—nearer geographically or in taste—will be firmer in their allegiance, and will desert to rival sellers only if the incentive is a strong one. Finally, the market may be imperfect because of preferences which do not originate in real differences between the products of each seller but in fancied differences or in habit. If I buy from Blank instead of from Blink, I may do so, not because experience has shown me that Blank's cigarettes, or petrol, or pills, are better than Blink's but from sheer force of habit, or

because, after reading the advertisements, I am biased in favour of Blank's, or because, without making a thorough test, I *believe* that Blank's goods are better suited to my purposes. I may also buy from Blank because he is a friend of mine, or because he buys from me, or because he went to the same school or goes to the same church. Blank has my goodwill, and I am part of his special market.

If the products of competing sellers are sharply differentiated from one another, any one competitor, by reducing its price slightly, will be able to draw custom from a wide field, and no firm will feel the loss of custom seriously enough to retaliate. We can then apply the ideas of marginal cost and marginal revenue in order to determine what price will yield maximum profits. Price will be above marginal cost and the ratio of the two will be governed by the elasticity of demand for the firm's product. If, however, the danger of retaliation exists, then this formula cannot be applied: we can calculate neither marginal revenue nor elasticity of demand until we know how large an increase in sales will follow a given cut in prices, and if we are uncertain about the train of events that will follow a cut in prices (e.g., whether or not there will be retaliation by other firms) then we cannot estimate precisely how sales will respond. We can, indeed we *must*, form some sort of judgment, but our judgment will be subject to a wide margin of error.

**Price-Competition and Service-Competition.**

When the market becomes imperfect, price ceases to be the sole consideration of consumers. A slight reduction in price by one seller, instead of drawing custom from all over the market, may pass completely unnoticed. A large reduction in price will certainly bring new customers; but if they come in large numbers from the clientèle of rival firms, the reduction in prices is likely to spread throughout the market until most of the new customers are tempted back to their original suppliers. Price as an instrument of competition, therefore, becomes either increasingly ineffective or increasingly unpopular. It is condemned as foolish if it is unsuccessful, and as cut-throat competition if it is successful. This is the point of view of business: the public, which enjoys the benefit of lower prices is not likely to feel itself injured. If no one took the lead in cutting prices how should we ever get things more cheaply, and how should inefficient firms ever be made to give way to their more efficient competitors?

The decline of competition in price has diverted competition into other channels. Price-cuts may be contrary to trade etiquette, but other devices of extracting custom are still sanctioned: advertising, improvements in quality or in service, the offer of a wider variety of goods and so on. All these devices have this in common: that they represent an actual *outlay* with a view to increasing sales, whereas the device of lowering prices represents a sacrifice in *revenue* with the same end in view. When competition reduces prices, the gain to consumers is unmistakeable: they save money which they would otherwise have had to spend. But when competition improves

quality and makes services more lavish, the gain is more debatable; consumers get more for their money, but they might get still more if they could have the old quality and service at a lower price and were free to spend the difference in price on other articles of more utility. To the man who can barely afford a third class fare in a Transatlantic liner, the competition which drives each steamship company to offer him a fast trip and the best of food is of less service than the competition which would bring down the fare. Competition through advertisement is of even more doubtful advantage. If it is purely informative, advertisement is very likely to render a service of real value; on the other hand, if it is purely percussive (i.e., striking home by sheer reiteration) the service rendered would hardly appear to be worth the cost.

Competition in services is wasteful in two ways. First, it provides us with goods of a better quality, and of a wider variety than we want, advertised more extensively and with more sales effort than we want. Second, it allows an excessive number of firms to survive; not only do the services undertaken by competing firms overlap (as in milk distribution), but also no firm is able to expand freely until it reaches optimum size (as it would if its competitors could not shelter within their special market). These wastes are sometimes referred to as wastes of competition. In fact, however, they arise, not because there is competition, but because there is not competition enough. If competition were perfect, both types of waste would disappear.

**Selling Costs.**

We considered at the beginning of this chapter a firm which was free to vary the price which it charged and the amount which it produced. We reached the conclusion that the firm would try to equate marginal cost and marginal revenue. But we took no account of a third variable in the policy of the firm—selling costs. Selling costs include all expenditure designed to create or increase the demand for a firm's products. Advertising expenditure is the most obvious example, but improvements in quality, special services, etc., have exactly the same effect. They modify the preferences of consumers by making the firm's products, in appearance or in fact, better suited to consumers' requirements, and so increase the firm's sales just as advertising does.

The existence of selling costs forces us to modify, but not to abandon, the marginal theory. The firm will push its sales costs, just as it pushes its production costs, up to the margin at which the last £1 spent on advertising, or on free gift schemes, or on any other kind of sales pressure, just repays itself by yielding a net increase of £1 in revenue (*net*, that is, after deducting the cost of producing the extra goods which are sold). Marginal sales cost, in other words, will tend to equal revenue from the last unit of sales outlay, just as marginal production cost tends to equal revenue from the last unit produced. With this reservation: that whereas the revenue from the last unit produced can occasionally be estimated with some confidence, the revenue from the last unit of sales outlay is almost altogether beyond

human conjecture. Who can predict what £1,000 spent on advertising will yield ?

We have seen that there are two margins—one of quantity produced, one of selling outlay. Two, however, is a gross understatement: there is a whole series of margins, cutting all the paths of policy open to the firm. The firm may consider, for example, a slight change in quality, or in finish, or in packaging, or in style. A change in any of these will involve an increase in outlay and in receipts, and the firm will move along the path of change until it reaches the margin at which a further move will add more to outlay than to receipts. *Any* marginal increase in outlay will, if maximum profits are being earned, involve an exactly equal increase in receipts, and marginal cost and associated marginal revenue will tend *everywhere* to equality.[1]

**2. Oligopoly.**

A perfect market is not by itself a guarantee of perfect competition. It is necessary also that there should be a large number of competing firms and that competition between them should not be limited by understandings or agreements. If there are only a few independent firms competing in a perfect market, we have a situation which is generally described as pure oligopoly. Each firm is producing such a large proportion of the commodity that a slight change in its output and sales has an appreciable effect on the market price. Each firm, therefore, will have an interest in restricting production for fear of forcing down prices; or alternatively, each firm will have an interest in refraining from cutting prices, for fear other firms retaliate. Village butchers, for example, may keep up the retail price of mutton while the wholesale price is falling; each knows that if he takes the lead the others will follow suit, leaving profits lower all round.

Under oligopoly it is clear that price will be higher than marginal cost of production. How much higher it is impossible to say. Each firm, faced with the danger of retaliation, must make up its mind how its competitors will behave if it reduces or raises its price; but no firm can puzzle this out without either a shrewd understanding of the psychology of its competitors or a long series of experiments in which these competitors are forced to show their hand. It is doubtful, however, whether, in practice, understanding and experimentation play much part. Oligopolists would seem to keep prices almost as high when independent of one another as when they are joined together in a cartel. In practice, too, oligopoly is rarely "pure." Where there are only two or three competitors they have generally their own special market and are in keen competition only

[1] It will not be true, of course, that *any* slight increase in costs should lead to an equal increase in receipts. The increase in costs must be such that *the best possible use* is made of the money in the direction in which it is expended. There may be a variety of changes in style which could be made at a cost. A large number will clearly not be worth while: some will be worth considering: and a true marginal change will be one which just barely pays for itself.

for " marginal " customers. If there are only two firms, for example, one will generally sell goods of high quality and the other standardised goods of lower quality.

**Conclusions.**

Under perfect competition price and marginal cost will tend to be equal, but where there are elements of monopoly, firms will be afraid to spoil their market and will maintain prices above marginal cost. This fear of spoiling the market will arise if the market is imperfect or if the number of competing firms is small. Market imperfection gives each firm an incentive to limit output to the requirements of its special market, while competition between a limited number of independent firms gives each an incentive to restrict output for fear of provoking retaliation.[1]

**Price and Marginal Cost : (2) The Industry.**

There are two sides to monopoly—high prices and high profits. So far we have dealt only with prices. We must now turn to a discussion of profits. This brings us to marginal cost in the second sense, i.e., average cost of production in a marginal firm.

An industry is generally made up of firms varying widely in profitability. Some firms earn high profits because of the superior judgment or organising ability of the management, others because their plant is well laid out and up to date, others because they have access to a large market or to cheap raw materials, and others again because of the skill or energy of the workers whom they employ. On the other hand, some firms earn comparatively low profits because they lack these advantages. In coal-mining, for example, there is a wide range of profits between one pit and another according to the depth of the coal seams from the surface ; the thickness of the seams ; the chemical composition of the coal ; the presence or absence of water or gas ; and a hundred other factors. Some pits rarely make losses while others rarely make profits.

[1] The phrase " spoiling the market " is applied above to a reduction in price which either provokes retaliation or is shared equally by established customers (who are willing to pay a higher price). The phrase can, however, be used in other senses. It can be applied, for example, to reductions in price which :

(*a*) by creating an expectation of a further fall, cause demand to fall off instead of stimulating it ;

(*b*) stimulate sales at the expenses of sales at a later date. Consumers or dealers may store their purchases and refrain from buying when prices go up again ;

(*c*) accustom the public to cheapness in price or quality and create opposition to any subsequent attempt to raise prices. What was intended as a temporary reduction in price may easily become a permanent one.

It is clear that the forces making against price-cuts are even stronger than might appear from the text. Yet price-cuts do occur. The reader will find it instructive to find reasons why this should be so.

The firms in the industry which we take to be marginal are those which, over a period of years, earn a profit just sufficient to induce them to remain in the industry. This minimum profit—which may be high or low—is really an element in the costs of these firms. A firm which has no prospect of earning a reasonable profit will ultimately close down, just as a man who sees no prospect of earning a wage commensurate with his abilities will ultimately seek employment elsewhere; the firm will have failed to cover its total costs.

Thus a marginal firm is one whose average costs (including normal profits) are just covered by price. If the price falls, marginal producers will be squeezed out and a new group of firms will find themselves on the margin of production. If the price rises, the margin will be pushed back as more firms find it worth while to come into the industry. The pressure of a change in demand and in price will cause the margin of production to shift inwards to firms whose costs are lower or outwards to include firms with higher costs of production. At a given level of demand the price will tend to equality with the average cost of production of the marginal producer. Other firms, in receipt of the same price, and producing at lower cost, will be able to earn a surplus profit—a profit higher than would suffice to keep them in the industry.

At any given moment, of course, price may fall below marginal cost. It is only over a period of years, averaging profits against losses, that the marginal producer will look for a certain minimum balance of profit. This makes it difficult and sometimes impossible to identify the marginal firms in an industry. We cannot always pick on the firms which are making the smallest profits or the biggest losses, for the firms which are making losses now may survive to make large profits in the future. They may, for example, be young and growing firms which have not yet solved their problems of organisation and marketing, but are on the way to a solution. Even if they are firms which are certain to go out of business sooner or later, we may still hesitate to call them marginal firms. We may know that at current prices there will be a constant succession of firms which will take their place with no more success—that it is a normal feature of the industry, in fact, that there should be some firms in it which cannot show a profit at all, even in good times.[1] Such firms are not marginal, but sub-marginal; their costs are higher than the costs of marginal firms, and will not be covered by price, even in the long run.

It must not be supposed that marginal firms are necessarily the least efficient firms. It may happen, for example, that some firms are efficient in manufacturing a number of different lines of product and that they are ready to abandon any one line whenever it is less remunerative than the others. A farmer may grow wheat, barley

[1] Take retailing, for example. In one American city where the average number of shops over a period of 22 years was 162, no less than 526 shops started business, and 354 dropped out.

and potatoes at low cost and yet be a marginal wheat farmer because the profit which he makes from his wheat is barely enough to keep him from giving up wheat growing in favour of barley or potato growing. He is on the *margin of transference* between wheat and other agricultural products.

Price, then, tends to be equal to the average cost of marginal producers and to exceed the average cost of other producers in proportion to any special advantages which they enjoy.

**Costs of Entry.**

So far we have assumed implicitly that the number of competing firms is large and that new firms can enter the industry freely; we have assumed also that there are some high-cost firms on the margin of production, and some firms with slightly higher costs, just beyond the margin, ready to enter the industry whenever prices rise. But what if potential competitors are under a handicap, so that they are attracted into the industry only when prices are *above* marginal cost? Marginal producers will then be making abnormal profits, while new firms that might produce at less cost and still make satisfactory profits will be debarred from competing. New firms will have to meet costs of entry as well as costs of production, while established firms, protected from competition, will have no costs of entry and will reap correspondingly high profits. These excessive profits, since they result from impediments to free competition, are essentially monopoly profits.

What are these impediments to free competition? Sometimes they are due to legal restrictions. The Elizabethan monopolists, for example, were able to buy complete immunity from competition by a cash payment to the Crown. The holder of a patent enjoys a statutory monopoly for sixteen years. Electricity supply companies can obtain a franchise from Parliament conferring on them a monopoly for a given area. The Traffic Commissioners can limit the number of licences issued to public service vehicles as they think fit, and can block competition from new transport companies very effectively. The British Broadcasting Corporation, the Bank of England, and the Agricultural Marketing Boards are all protected by the State in one way or another from competition.

When legal restraints are absolute, so that there is full monopoly, prices might seem free to rise without limit. In fact, however, there is a double check.[1] The monopolist will not wish to lose his monopoly by too obvious abuse of his powers; what the law has given, the law, under pressure of public opinion, can take back.[2] Secondly, excessive prices will lose not only goodwill but customers; there is a limit to what people will pay, as well as to what they will tolerate. The more elastic the demand, the sooner will this limit be reached, and the greater, therefore, will be the resistance to a rise in prices.

Legal restrictions make costs of entry infinite. So also does complete monopoly of raw materials; the International Nickel Company, for example, has no need to fear competition from other nickel

producers in fixing prices. But when the monopoly of raw materials is a local one only (e.g., the Aluminium Company's monopoly of bauxite deposits in the United States) then costs of entry become finite and measurable. They will, in fact, be equal to the cost of transport of alternative supplies from the nearest available point. Potential competitors are handicapped by distance from the local monopolist's market so that if their costs are equal to the monopolist's and there are no economies of large-scale production, the monopolist can safely augment his price by the cost of transport from the nearest alternative source of supply. He will then make a surplus profit on each unit of output equal to this addition to his price. The surplus (or monopoly) profit will be equal to the cost of entry into his market or—what comes to the same—to the gap between price and marginal cost.

Limitation of entry may also arise when existing firms have exclusive control over the channels of distribution. They may be able, for example, to black-list new firms and cut them off from the normal retail outlets. The mere fact, too, that existing firms have already built up a business connection gives them a strategic advantage over new firms, and may enable them to smother competitors who, although more efficient, have difficulty in gaining a foothold in the market. This strategic advantage will be particularly great where heavy capital outlay either on advertising or on fixed plant is necessary before production can be started. The newcomer must plunge boldly in instead of feeling his way. Hence he will run a heavy risk of failure without the chance to cut his losses. If the market is narrow he will have to face the additional risk of bringing down prices to levels unattractive to himself and unprofitable also to his competitors. One village cinema, for example, may find it easy to make a large profit. But a second cinema may introduce overcapacity and leave neither cinema with a satisfactory profit. Thus a firm which has the advantage of priority in a narrow market—particularly in industries using large amounts of fixed capital, like the steel or the railway industry—may continue to earn profits higher than a new firm of equal efficiency would consider adequate. There may be an appearance of free competition. But costs of entry, due to the narrowness of the market, the initial risks on a heavy capital investment in face of competition, and the hold of the established concern both on supplies of labour and materials and on the market, will combine to shelter the existing firm from competition and protect it in the enjoyment of a monopoly profit.

**Summary.**—The tendency of competition is to reduce prices to equality with cost—with the marginal cost of each firm, and with the average cost of a marginal firm. The tendency of monopoly is to raise prices above cost and profits above normal. In the modern world we normally meet with a mixture of competition and monopoly. There are substitutes, but not perfect substitutes, for each firm's products, and there is generally a threat of competition from newcomers without complete freedom of entry.

One final point. Although the tendency of monopoly is to raise prices above costs, this does not imply that prices under monopoly are higher than under competition. The costs of the monopolist may be so far below the costs of the competing firms which might replace him that his price also is lower than theirs would be.

## CHAPTER 15

## SUPPLY

## (2) COST IN RELATION TO OUTPUT

THE fact that cost of production varies with output is constantly overlooked. People speak of *the* cost of production of an article as though there were only one cost, irrespective of the *amount* of the article produced. But it is well known that in some industries (e.g., the motor industry) costs fall when a substantial increase in output takes place, while in other industries (e.g., wheat farming) costs rise when output is increased. Why is this?

The reason why costs may fall as an industry grows bigger has already been explained. If the industry is to expand, either the constituent firms must produce on a larger scale or new firms must grow up alongside the old. The first method of expansion may very well result in internal economies, the second in external economies. Economies of either type will reduce the average cost of production.

The reason why costs may rise is more obscure. We have seen that economies of scale are limited, so that no firm can grow beyond a certain size without loss of efficiency. Costs *within each firm* must ultimately begin to rise as output expands. But this does not mean that sooner or later additional supplies of the commodity will only be available at increasing cost. For if all firms are alike, and there is free competition, no firm will ever pass the optimum size. Whenever existing firms reach that size, an unlimited number of new firms, of the same size and efficiency, will be ready to enter the industry and provide, at constant cost, for any increase in demand, however great.

In the last chapter we saw that firms are not, in fact, alike and that freedom of entry into an industry is generally limited. Here we have two possible explanations of rising costs. As an industry expands it attracts new firms which have either progressively higher costs of production or progressively greater difficulty in gaining a foothold. The second explanation, which will not be pursued here, turns on an increasing degree of monopoly; the first, which is by far the more important, turns on the scarcity and heterogeneity of productive resources.

The first explanation has so far been expressed in terms of firms; when prices rise, and the margin of production is thrust back, the

new firms entering the industry are less favourably situated than the old. But it is not really the new *firms* that are more costly; it is the new *factors of production* attracted into the industry. If the cost of coal rises when the output is increased, it may be because sub-marginal coal deposits have to be mined, or because the new seams are worked by less experienced miners, or because the new pits are less ably managed. It makes no difference which factor is responsible for the rise in costs, and it makes no difference whether that factor is employed by new or by existing firms. The important point is that an *industry* can obtain command over additional labour, land, raw materials, managerial ability, etc., only by the offer of increasingly generous terms: by drawing in units of lower efficiency or with a higher supply price.

The industry must either use resources that were formerly left out of employment altogether, or bid against other industries for resources in their employment. Farmers, for example, can grow more wheat either on sub-marginal land or on land which would otherwise have been sown to oats or barley. Motor manufacturers can obtain more steel either by inducing the steel industry to bring more furnaces into production or by leaving less steel for shipbuilding, tin-plate manufacture, etc. Expansion, in short, means the use of a larger share of the limited supply of productive resources, and these resources will be released by other industries only when they find themselves unable to bid high enough to retain them. Additional resources, therefore, can be commandeered by an expanding industry only at increasing cost.

**The Meaning of Cost.**

This line of reasoning leads us to a new view of cost. The cost of producing any commodity is fundamentally the cost of *detaining* productive resources so as to make their services available to the industry producing the commodity *rather than* to some other industry. Now the cost of detaining resources is equal to the value of these resources for other purposes, plus the cost of transferring them from their existing employment to the most attractive alternative employment.[1] Cost, in other words, measures the pull of competing attractions. If the cost of any factor is not covered, that factor will move to some other industry, and the value of what it could produce in that industry is its cost in its present use. Cost, therefore, must always be considered in terms of alternatives. Cost measures what we could produce instead—what we forgo by using productive resources for one purpose rather than for another. Alternatively, it measures what we could save by not using the resources at all. If, for example, labour were employed for seven hours a day instead of eight, there would be a fall in the output of goods and services,

[1] It may happen that other industries are paying resources less than their full value, so that the force of attraction is diminished. For the sake of simplicity, this possibility is ignored.

but labour would enjoy more leisure and would be relieved from disagreeable duties. The sacrifice of leisure and the additional irksomeness involved in working for eight hours measures the cost of the additional goods produced in the final hour.

Two illustrations may help to make this view of cost clearer. When I give a friend the use of my car for the day "at cost," how is he to decide how much to pay me? If he consults my garageman he may be told that £1 is a fair amount, since the garageman calculates that he could not afford to keep cars for hire for less than this. My friend may then point out that I already own the car, that I have to provide garage accommodation for it anyhow, and that if he had not borrowed it, it would not have been used at all. Provided the car is returned undamaged, I am no worse off by letting my friend have it free. The cost, therefore, is nil, because the *alternative* of letting the car stand idle would be just as expensive. A second illustration can be taken from the argument that a programme of road-building at the bottom of a slump would be too costly. We are faced with the choice between supporting large numbers of men in unemployment and using their services to build roads that will be of permanent value to the nation. The social cost of the roads is the value of the unemployment which we get in place of them, and it does not require much imagination to see that this may come to a fairly large *negative* total.

**Cost and Rent.**

It must not be supposed that every unit of resources is paid no more than its cost. It is, in fact, only the marginal units which earn just enough to keep them from moving to some other industry. Other units, which have a greater bias towards their present employment, will earn a surplus above their "transfer cost." This surplus, which is exactly similar to the surplus profit earned by intra-marginal firms, is called "rent." Thus the earnings of any unit of resources are made up of two parts—"transfer cost" and "rent." The "rent" element in earnings will be high, where a factor is specially proficient in its present use, or suited to it temperamentally, or disinclined to move elsewhere. The idea of rent is one to which we will return in later chapters.

**Scarcity, Heterogeneity, Mobility.**

We have seen that rising costs originate in the scarcity and heterogeneity of the factors of production. The tendency for costs to rise in any industry is the more pronounced the scarcer are the factors which it employs (i.e., the larger the proportion of the total stock engaged in the industry) and the more heterogeneous these factors are (i.e., the poorer the substitutes on which it can draw). In an industry like building, which employs a very large proportion of the bricklayers, masons, slaters, etc., in the country, there is a strong tendency towards rising costs; whereas in radio manufacture, which employs a comparatively small proportion of electricians, the

tendency towards rising costs is almost negligible. Similarly, since the difficulty of finding more land suited to wheat-growing is greater than the difficulty of finding more land suited to dairy-farming, the cost of wheat tends to rise more steeply than the cost of milk. There is a third factor on which the tendency to rising costs depends—the mobility of productive resources. The greater the ease with which the factors of production can move from one industry to another, the less rapidly will costs rise. If, for example, labour is highly adaptable, or if most industries use a somewhat similar technique, then the cost of expanding any one industry will be much less than it would be if every worker were given a long and narrow training which fitted him for one job and one job only.

**Relative Scarcity.**

When an industry expands it may be specially difficult to obtain an adequate supply of some one factor. If no increase is possible, and there are no substitutes for the factor, the industry will be unable to expand further. But if there is some alternative technique which eliminates—in part or altogether—the use of the scarce factor, then expansion will be able to proceed at (presumably) rather higher cost. Resort to such alternative techniques, in order to meet a relative scarcity of some one factor, takes place every day. The problems raised by these changes in technique call for special consideration.

We start from two sets of facts—that great changes take place from time to time in the scarcity of one factor of production in relation to the others, and that corresponding changes take place in the use made of that factor relatively to the other. In some ages and countries labour is abundant and capital is scarce ; methods of production, therefore, are such as to economise capital and use plenty of labour. In other ages and countries it is capital which is abundant and labour which has to be economised. The same services are rendered by combining the factors of production in very different proportions. Harvesting in China, where capital is scarce, is done with the scythe ; in Canada, where labour is scarce, it is done with the combine harvester. In passenger transport there are endless possibilities of variation—from the labour-consuming rickshaw to the capital-consuming Transatlantic liner. The motor omnibus, the taxi, the railway sleeping car, the compartment for 8 hommes, 40 chevaux—all provide transport services with varying combinations of labour and capital.

When we say that the factors of production can be combined in different proportions we are really saying that they are substitutes for one another. We can produce the same quantity of product by using more capital and less labour (substituting capital for labour) or by using less capital and more labour (substituting labour for capital). The ease with which we can change the proportions by substituting one factor for another varies from industry to industry. In some industries substitution is practically impossible ; the technique of the industry is inflexible and dictates the use of productive agents

in a fixed, or almost fixed, proportion. In the building of brick cottages, for example, both the materials required and the number of bricklayers, plumbers, carpenters, slaters, etc., are practically fixed once the plans are drawn.[1] In most industries, however, technique is fairly elastic, and a factor which has become specially cheap can be substituted for factors which remain dear. In the steel industry, for example, steel scrap and pig-iron can be substituted for one another in the making of steel whenever one of them becomes relatively scarce and expensive.

There are two general principles which govern the substitution of one factor for another. The first is that no one factor is a perfect substitute for another. This follows from the definition of a factor of production; if factors could be freely substituted for one another in all uses they would all belong to a single group of factors. The second principle, which is based on experience, is that substitution becomes progressively more difficult the more of one factor we try to substitute for another. We can easily do without a little of one factor and use a little of another instead; but we cannot easily do without a great deal of one factor and replace it by another. A small change in the pig-iron-scrap ratio makes little difference, but a large change creates great technical difficulties. The two principles, taken together, give us the "law" of diminishing returns to the factors of production. This is simply a statement that, sooner or later, other things remaining the same, the combination of an increasing number of units of one factor with a given number of units of other factors must lead to a less than proportionate increase in output. The total product increases, but it does not increase so rapidly as the variable factor. The reason for this is just that some factors are not increased at the same time as the variable factor, and that increased supplies of the variable factor cannot entirely make up for this deficiency unless the variable factor and the fixed factors are perfect substitutes for one another.

The law is illustrated in Table III. The first three columns of the table are hypothetical and provide the data from which the last two columns are calculated. Column 4 is obtained by dividing the total product by the number of units of A in use. If column 4 is increasing as output increases, the increase in output is more than in proportion to the increase in A; and if column 4 is decreasing, the increase in output is less than in proportion to the increase in A. In the first stage we have increasing, and in the second decreasing, returns. A third stage is also possible, intermediate between these two, at which output increases exactly in proportion to the increase in A, so that for a time column 4 remains constant. At this stage we have constant returns.

According to the law of diminishing returns we must ultimately reach the stage at which output increases less rapidly than units of A.

[1] Even here there are possibilities of variation, e g., by substituting more skilled for less skilled men, or high quality for low quality materials.

But is there any reason why we need pass through the earlier stages of increasing or constant return? If, for example, we can produce 1,530 units of product with six units of A and ten of B, why isn't it possible to produce half as much (i.e., 765 units of product) with three units of A and five units of B, instead of the mere 600 units which our table gives as the joint product of three units of A and ten units of B? Why can't we save five units of B and still turn out 165 units more? Isn't there an obvious waste in using *any* of the first five combinations, whatever the cost of the factors or the price of their product? Is there any reason, other than muddle-headedness, why these combinations should ever be adopted in the real world? The answer, to take an example we have just given, is that it may not be possible to use only five units of B, or find a market for 1,530 units of product. If B is indivisible (i.e., if it

TABLE III.

| Units of A | Units of B | Units of Product | Average Product per Unit of A | Average Cost of Production when A=£20, B=£10 |
|---|---|---|---|---|
| 1 | 10 | 50 | 50 | 2·40 |
| 2 | 10 | 200 | 100 | ·70 |
| 3 | 10 | 600 | 200 | ·27 |
| 4 | 10 | 960 | 240 | ·19 |
| 5 | 10 | 1,270 | 254 | ·16 |
| 6 | 10 | 1,530 | 255 | ·144 |
| 7 | 10 | 1,750 | 250 | ·137 |
| 8 | 10 | 1,920 | 240 | ·136 |
| 9 | 10 | 2,070 | 230 | ·135 |
| 10 | 10 | 2,200 | 220 | ·136 |

cannot be duplicated on a smaller scale so that a smaller number of units can always be taken if desired) it may be necessary to use one of the earlier combinations of A and B so as to produce a limited quantity of product; and, secondly, it may be necessary to limit output because the market is narrow. A small motor-factory, for example, may have half the capital of a large one and, although employing more than half as many men, may turn out less than half as many cars. It is not of much use to recommend the small factory to adopt the methods of the large one and turn itself into a kind of half-scale replica, for it is only factories which are of a certain minimum size that can avail themselves of the methods of the large factory. Nor is it worth while to recommend the small factory to double its output and become a whole-scale replica of the large factory, for the increased output may not find a market. In

other words, where there are economies of scale and market resistances to expansion we will find increasing returns operating, and continuing to operate.

Suppose now that units of A cost £20 and units of B £10. The tendencies to increasing and decreasing return are then converted into tendencies to decreasing and increasing cost. The average cost of production falls until output is at 2,070 units and rises for larger outputs. It will be observed that the output at which average cost begins to increase is considerably greater than the output at which returns begin to diminish. This is due to the spreading of the fixed cost of £100's worth of B over each output. Cost per unit of A is at a minimum when the return per unit of A is at a maximum. But cost per unit of A *plus* B continues to fall for some time. If a large amount of capital has to be sunk in a railway, the total cost per passenger will go on falling long after each additional train has begun to carry a smaller and smaller number of extra passengers.

The calculations of Table III are intended to show how the substitution of one factor for another is, after a certain point, increasingly ineffective. It is assumed that only a strictly limited supply of one factor (B) is available, and that there is an unlimited supply of another factor (A). This assumption is obviously an extreme one. But if we remove it, and suppose instead that the supply of B is *more limited* than the supply of A, or that B is *relatively* scarcer than A, the principle illustrated in Table III is unaffected. As output expands there will be a shortage of both A and B, and the cost of both will increase. But since B will be relatively scarcer and more expensive than A, it will pay to use a larger proportion of A and to economise B. The change in the proportion of A to B, whatever the absolute amounts of A and B used, will bring the law of diminishing returns into play, exactly as in Table III. The substitution of the relatively cheaper for the relatively dearer factor can keep down cost of production only within limits.

Thus if all the factors employed in an industry are equally variable there will be no reason—apart from economies of scale—why one should be substituted for the other as output increases. The two will be combined in the proportion which makes costs a minimum, and if output has to be doubled, twice as much of each—again apart from economies of scale—will be used. If it is open to every industry, for example, to draw freely on land, labour and capital at current prices, each industry can double its output without increase in cost; there will simply be two factories for every one there was before. So long as there is no reason to economise one factor rather than another, and so long as both are available in abundance at a constant price, costs cannot rise. If the second condition is not fulfilled, costs will rise because both factors are costing more as the industry seeks to attract increased supplies of each of them. That is, costs will rise because of scarcity. If the first condition is not fulfilled, if pressure on one factor is greater than pressure on the other, so that the only units of the first factor that can be recruited are costly or

second-rate, while the second factor can still be obtained in comparative abundance, then there will be good reason to substitute the second factor for the first. If such substitution were impossible and the factors had to be combined in a fixed proportion, costs would rise steeply. If substitution were perfect, and the second factor obtainable without limit at a fixed price, costs would not rise at all. In practice, as we have seen, the law of diminishing returns makes substitution imperfect and costs rise none the less. Thus it is because industries are forced to make do with the factors of which they *can* make increased use, meeting a deficiency of other factors as best they can, that diminishing returns come into play. Using a larger proportion of the factor which can be most easily increased is simply a method of economising the other factors and avoiding the consequences of their scarcity.

**Conclusion.**—Economies of scale, and economies of scale alone, make costs fall as output increases. If there were no economies of scale, increasing returns would disappear; and if increasing returns disappeared, decreasing costs would go too. Scarcity, combined with the law of diminishing returns, makes costs rise with output. Scarcity can be partially circumvented by substitution of one factor for another, but this substitution is limited by the tendency towards diminishing returns.

The laws of increasing and decreasing returns (and the corresponding laws of decreasing and increasing cost) are often cited as if they were in some way parallel to one another. But it is now clear that they are quite distinct. Scarcity, in which the law of decreasing returns originates, has no connection with economies of scale which give rise to increasing returns. Nor are the industries in which the two laws operate distinct from one another. Each industry has its bottlenecks —points at which the scarcity of some factor or factors presses heavily—and each industry has its economies of scale. In some industries economies of scale predominate and in others they are comparatively insignificant. In the first group costs will be falling, and in the second they will be rising, as output increases. But the laws of decreasing and increasing returns operate simultaneously in each; it is only *on balance* that we can draw a line of division between them.

It is often said, for example, that agriculture is subject to decreasing, and industry to increasing, returns. But agriculture is full of examples of economies of scale—through the use of the combine-harvester, tractors, milking machines, and so on. It happens that the scale of production at which these economies can be realised is very small, while the market for agricultural products is very large. It is rare, therefore, to find any branch of farming working under conditions of increasing return. Scarcity, on the other hand, makes itself felt in farming very markedly: There is a scarcity of fertile land; so that, if the output of all farm products has to be increased, farmers must either bring into cultivation sub-marginal land, or cultivate the existing area more intensively by applying to it an increased amount

of labour and capital. The first expedient raises land farming costs because sub-marginal land is inferior in fertility or in situation to the land already under cultivation. The second expedient raises farming costs because, if the proportion of labour and capital to land is increased, the tendency to diminishing returns comes into play. Each successive application of labour and capital yields a diminishing increment of product. If the output of only one farm product, e.g., wheat, has to be increased, scarcity will again force up costs. Land below the margin—this time the margin of transference—will have to be brought under wheat, and/or existing wheat land will have to be cultivated more intensively. The first expedient means paying a higher rent or renting inferior wheat land; the second brings decreasing returns into operation. If agricultural labour is also scarce, and the channels through which capital can be borrowed or credit obtained are narrow, then the rise in farming costs will be all the steeper.

In industry, on the other hand, the scale of production at which fresh economies cease to be made is enormous, and the market is comparatively narrow. Thus it is rare for an industry to be so placed that it could not more than double its output with the use of twice as much land, labour and capital. The average industry is producing under conditions of increasing return. The pressure of scarcity, too, is less severe. The capacity of the industry can be increased by building plant exactly similar to the plant already in use. More workers can be found and trained much more easily and quickly than agricultural workers. Capital can be obtained in almost unlimited quantities. Thus costs of production rise only when output has to be increased very hastily, and when scarcity, therefore, is at its greatest. In the long run, costs in industry generally fall when output is increased.

## CHAPTER 16

## SUPPLY

## (3) COST IN RELATION TO TIME

WHEN there is a change in the conditions of demand or supply, the full effects of such a change are rarely felt immediately. Consumers may go on buying the same goods for some time after prices have risen and only gradually transfer their expenditure to other goods; producers may continue in business for years at prices which yield them a poor, or even a negative, return on their capital, and may only give up when their plant is completely worn out. Generally demand adjusts itself to changes in price much more rapidly than supply, and in the short run, therefore, dominates the course of prices; in the long run, however, supply becomes progressively more important as prices are pulled into line with cost.

A period sufficiently long to allow a full response of demand and supply to a change in prices is called "the long period." A period so short that there is no time for the "capacity" of an industry (its fixed plant and organisation) to alter appreciably is known as "the short period." But there are, of course, a whole series of periods, of increasing lengths, in which successively closer approximations are made to a full adjustment of demand and supply. This full adjustment, in which there is no further tendency for price to change, is described as "equilibrium." In the long period, therefore, price will always reach its equilibrium level. Equilibrium is something which never exists, but is always coming into existence; price never has time to adjust itself finally to one change in demand or supply before another change interrupts progress towards equilibrium. The fact that a position of long-period equilibrium is never reached in the real world does not mean that the analysis of the last few chapters (which is almost exclusively in terms of long-period equilibrium) is quite futile. There are always long-period tendencies at work in each successive short period through which we live; and long-period equilibrium is simply the logical conclusion of these tendencies—their end-product, other things remaining the same. Moreover, if we understand what governs prices in the long run, we shall find it comparatively easy to apply the same line of reasoning to shorter periods.

**The Very Short Period.**

In the very short run—say, in the day's marketing—supply can be altered only by drawing on or adding to stocks. If the commodity is highly perishable there will be no stocks, and the influence of supply on price will be nil. In Covent Garden, for example, the supply of strawberries on sale every day will be fixed by the deliveries arriving in the morning, and will not respond to a rise or fall in prices during the day. The price of strawberries will be governed by the demand, in the sense that the price must fall to the level at which the whole supply will be bought. The demand is likely to be fairly elastic, since many buyers will be able to hold off until the following day if prices are abnormally high, and will be able to increase their purchases if prices are abnormally low. There will probably also be some minimum price at which sellers prefer to keep back their strawberries for disposal to jam manufacturers, and at this "reserve price" supply will become extremely elastic since there will be no limit to the surplus from one day's sales that can be made into jam.

A similar example, covering a longer period of time, might be taken from potato-growing, where no increase can be made to the supply for almost a year after the crop has been harvested. The price is driven to the level at which the fixed supply can be rationed out amongst consumers so as to leave no surplus unsold. But amongst those consumers are the potato-growers themselves. For if prices fall too low they will feed their potatoes to stock instead of selling them for human consumption. Like the strawberry growers, they have a reserve price at which they withhold supplies from the

market and below which, therefore, the market price cannot fall. In exactly the same way the price of commodities which are not perishable and can be stored cannot, in the short period, fall below the price which holders of the commodity expect to receive if they refuse to sell until later. This anticipated price (less carrying costs such as storage and interest charges) is their reserve price at which, so to speak, they enter the market on their own account.

**The Short Period.**

In the short period, supply can be altered through a decision on the part of any firm either to make a marginal change in its output, or to discontinue production for the time being. The first decision will hinge on the relationship between marginal cost and revenue; the second on the relationship between total cost and revenue. But neither the marginal nor total cost of producing an article is the same in the short period as it is in the long. Suppose, for example, that I own a season ticket between Cambridge and London, and I stop to ask myself how much a journey to London is costing me. The answer, until the season ticket expires, is "nothing at all." I have paid in advance for as many journeys as I care to make. Suppose now that I am wondering whether I ought to renew my season ticket. Then the cost of a journey to London becomes something positive and calculable. It is equal to what I can save by not travelling. If I mean to make only a few journeys, for example, then the cost of each is clearly equal to the ordinary return fare. In the short period while my season ticket is unexpired, the cost of a journey is nil; in the long period, it may be as much as the ordinary return fare. In the same way the cost of using a piece of machinery, once it is installed, is practically nil, but the cost of using the machinery if it has to be hired every time it is used is by no means negligible.

The fact is that cost is not the same in the short period as it is in the long, because the alternatives open to us are not the same. In the long run we have the alternatives of *not* renewing our season tickets or the machinery which we have installed; in the short run, this alternative does not exist. A cost can only be a cost if it is something that can be avoided. But we cannot, in the short period, avoid the cost either of season tickets or of machinery; the cost of both has already been met, and since "bygones are forever bygones" we cannot avoid the cost *now*. If the machinery depreciates more rapidly when in use than when it is idle, the *extra* depreciation can be avoided by not using the machinery and does, therefore, constitute a true cost, even in the short period.

It follows from this that costs which are fixed in the short period, and do not vary with output, are not really short-period costs at all. For example, firms cannot, in the short period, avoid payment of rates and taxes or debenture interest. Nor can they alter their fixed plant and organisation and so avoid the cost of depreciation and obsolescence. These costs are fixed, whatever the level of output; they have to be met even if nothing is being produced. There is,

so to speak, a cost of producing nothing. Now the cost of producing a given output is, as we have seen, what we could save by not producing it—what we could save by producing nothing. But the alternative of producing nothing may still involve us in expense. That expense, therefore, forms no part of the cost of what we do produce. It is only the additional *avoidable* outlay which is incurred that enters into the short-period cost of production. The cost of producing nothing is, for all practical purposes, the same thing as "overhead" or "supplementary" cost.[1] Overhead costs, therefore, are fixed independently of output and have, or should have, no bearing on the questions: What output, if any, is most profitable now? What is the highest price that I can charge and the lowest that I can accept? In the short period it will pay to produce goods even if they make no contribution whatever to overhead costs and fetch no more than their prime or avoidable costs; and in deciding what price to charge or to accept for his goods, a manufacturer will be guided more by his idea of what the market will bear than by the allowance which he thinks he must make for overheads. The plea that he must cover his overheads may be useful in mollifying his customers, and the suggested allowance for overheads may enable him to judge what price he can safely charge without inviting fresh competition from new firms. But it will be information on *marginal* costs—the cost of a few units more or a few units less—rather than on *average* prime and overhead costs that will be of most service to him so long as he remains in production. The allocation of overhead costs between different units of output (and also, as we shall see, between different lines of production) is, from the economic point of view, an irrelevant piece of ritual. In the long run, of course, overhead costs must be covered or manufacturers will stop producing. But they need not be covered consistently. Normally, some surplus over prime or variable costs will be earned. Sometimes this surplus will rise above and sometimes fall below overhead costs. It will fluctuate with the state of demand, increasing when an active demand allows prices to be raised, and decreasing when demand falls away and prices have to be reduced. The course of prices will govern the allowance that can be made for overheads. In the long run, however, the position is reversed—the allowance that has to be made for overheads must ultimately influence prices. How does this come about?

The answer is that in the long period overheads are true costs. Outlay on overheads can be avoided by going out of business, or by closing down plant, which is idle and expensive to maintain, or by failing to repair and renew buildings and plant. In the short periods

[1] An alternative (and less novel) definition of overhead cost may be simpler to understand. Overhead cost can be taken to mean the difference between total cost and the cost of those factors of production (e.g., labour and materials) which vary with output in the short period. This alternative does not require any considerable modification of the statements made above.

these alternatives do not exist. A business will be willing to spend money in order to remain a going concern producing nothing; it will be willing to meet the cost of producing nothing. It is only in the long run, when it has no incentive to remain a going concern, that it will prefer to produce nothing for nothing instead of at considerable expense. Similarly it will meet the cost of maintaining idle plant so long as there is a prospect that this cost of maintenance—not the original cost of the plant—will be recouped. The problem of replacements and renewals introduces even more elasticity for there will be some plant that falls due for renewal every year, perhaps even every month—and if prices are not high enough to justify renewal, output will immediately fall. What was a fixed cost in the lifetime of the old machine becomes a variable cost when the purchase of a new one is being considered, and all variable costs necessarily influence supply and price.[1]

We can now return to the two sets of decisions by which supply is altered in the short period. A marginal change in output may cost either more or less in the short period than in the long. A marginal reduction, for example, will save only the cost of labour and materials in the short period, while in the long period it will save part of the overhead cost of machinery and plant. On the other hand, a marginal expansion in output may, if the firm is working to capacity, be much more costly in the short period when there is no time to instal additional machinery or train additional men, and the firm has to make shift by overdriving its machinery and overworking its employees. Especially if there has to be more overtime, with correspondingly high rates of pay, the short period cost of a marginal expansion in output may be far in excess of the cost of a similar expansion over a long period.

The complications introduced by time make the problem of pricing an exceptionally delicate one. A firm has to weigh cost now and in the future against revenue now and in the future, trying to ensure that no marginal change in output will yield it a greater profit now without damage to profits in the future. It must hesitate, for example, to refuse a rush order from an important customer even when marginal cost is high—higher than price—if it fears that refusal will prejudice its sales in future years. Equally it must abstain from methods of production which leave a balance of profit now only by ruining all chance of producing at a low cost in time to come.

The second set of decisions—to shut down temporarily or open up again—depends upon the relationship between total cost and total revenue; in other words, if we confine ourselves to the decision to close down, upon how big a loss firms are making. But the *amount*

[1] It follows that when we speak of overhead cost we ought to specify some period of time during which cost remains fixed. Overhead cost is the cost of producing nothing in some given interval of time and in a given state of expectation. This interval of time will generally be one in which replacements of plant are negligible.

of the loss will not be the sole factor. We have also to ask, " *In what sense* is a loss being made ? " If, for example, the firm is making a loss on total costs (including overheads) it is not likely to stop producing. If it cannot cover its debenture interest it may, of course, be forced into bankruptcy. But since a prime profit is being made (i.e., since the firm is making a profit over prime cost) it will generally pay the creditors to allow production to go on. It is better to have some return than none at all.

If a prime loss is being made, the action of the firm will depend upon three main factors :—its financial strength, the view which it takes of future prospects, and the cost of closing down temporarily and opening up again.

(1) The *ability* of a firm to make a prime loss will depend on its ability either to raise money on loan or to draw on reserves. No firm can go on disbursing more each week than it receives in the form of sales proceeds without running into debt.

(2) The *willingness* of a firm to incur debt (assuming that lenders can be found) will depend upon its faith in such a revival in demand as will allow the debt to be repaid. A loss on prime costs is a species of investment on which a return is expected at a later date when the firm hopes to profit from having maintained its plant and organisation in good condition.

(3) The *wisdom* of continuing to make a prime loss will depend upon the costliness of the most attractive alternative. It may be cheaper, for example, to shut down temporarily, putting the plant on a care and maintenance basis. But this is a course which firms are reluctant to take, since it may mean the loss of business connections, disorganisation of staff, and the dispersion of a reliable and carefully selected body of workers, trained in the ways of the firm. Frequently, therefore, firms prefer to work on a skeleton output in order to maintain contact with markets, staff and workpeople. Where shut-down and reopening costs, broadly defined, are high, firms will prefer to go on producing at correspondingly high prime losses.

In the short period, as in the long, there will be some firms that are marginal and others that are intra-marginal. Marginal firms will be on the verge of closing down or opening up, and a small change in price will be sufficient to turn the scale. The average short-period cost of such a firm can be calculated by deducting from its total costs the alternative cost of closing down and producing nothing and averaging the residue over the firm's output. Below this average cost, price will not fall. Monopolistic influences may, however, succeed in maintaining prices above the average short-period cost of even the least favourably situated firm.

**Summary.**—A change in price will have immediate and delayed reactions on supply. These reactions will be governed at each stage by the alternatives open to producers, and the range of alternatives will widen with the passage of time. Supply is more elastic the longer the period which we have in view.

This is particularly true of a *reduction* in supply. What is saved by contracting output in the short period is small in comparison with what can ultimately be saved (e.g., by refraining from plant renewals). Especially when overheads are high, the adjustment of an industry to a *fall* in price is difficult and protracted. On the other hand, an industry can generally expand its capacity quickly to meet a *rise* in price because short-period expedients for increasing output tend to be much more costly than the long-period expedient of installing more plant.

## CHAPTER 17

## INTER-RELATIONSHIPS OF SUPPLY AND DEMAND

THE effect of a change in price is never confined to a single product; there are repercussions on the prices of other products, linked with the first either in supply or in demand. These repercussions can be classified under five headings:—

**1. Joint Supply.**

There are some commodities which can only be produced in association with other commodities; familiar examples, are wheat and straw, mutton and wool, gas and coke. These commodities are said to be in joint supply. Commodities are in joint supply whenever one commodity is a by-product of another, so that it is impossible to increase the supply of one without simultaneously increasing the supply of the other. If there is an increase in the demand for one joint product, therefore, the supply of *both* will increase, and, since there will be a larger supply of the second product to dispose of, its price will tend to fall. Similarly, a fall in the demand for one joint product will tend to increase the price of the other.

**2. Joint Demand.**

Some commodities are wanted in association with other commodities, and when the demand for one increases, the demand for the complementary commodity increases simultaneously. For example, there is a joint demand for tea and sugar, collars and ties, bacon and eggs, whisky and soda. If any one of these becomes cheaper because of an increase in the supply, both joint products will be in greater demand, and the price of the second, which is no more plentiful, will tend to rise. Similarly, if one joint product becomes scarcer, the price of the other will tend to fall.

**3. Composite Supply.**

Commodities in composite supply are substitutes for one another. Butter and margarine, for example, form a composite supply for the satisfaction of one want; tea, coffee, cocoa, etc., for the satisfaction

of another. If, of two commodities in composite supply, one becomes cheaper, the demand for the substitute commodity will clearly be reduced, and its price will tend to fall, too.

**4. Composite Demand.**

Commodities are in composite demand when they are wanted for several different purposes. Electricity, for example, is in composite demand for lighting, heating and cooking; for electric razors, clocks and wireless sets; and for driving electric motors in many different industries. If more electricity is needed for lighting, there will be less available for heating and cooking; an increased demand in one use will put pressure on the supply in alternative uses, and prices will tend to rise all round. In the same way, labour is in composite demand for all the industries in which it can be employed. If, then, more workers are urgently needed in one set of industries (e.g., the armament industries), the scarcity of labour in these industries will communicate itself to all industries employing workers of the type in demand, and wages will go up all round—in the industries losing workers as well as in the industries gaining them.

**5. Derived Demand.**

The demand for some commodities *derives from* the demand for other commodities. No one wants crude steel for its own sake; it is wanted for the manufacture of steel products. The demand for steel is a derived demand; yet from it, in turn, other demands derive—the demand for iron ore, coal, scrap iron and manganese. It is only when we reach the ultimate factors of production that the chain of derived demand comes to an end. At the other end of the chain are the final products of industry ready for sale to consumers. The demand for the services of the factors of production, therefore, is derived from the demand of the consumers of finished goods.

These five relations sum up the inter-connections of different industries and different markets. One product is tied to others in the most unexpected ways, and without a vivid sense of these obscure ties we cannot lay claim to a real understanding of the theory of value. Consider, for example, the ramifications of a programme of rearmament. More steel has to be produced. Hence there will be a larger output of furnace slag—to take one by-product out of many—and the price of cement and of fertilisers, both of which can be made from slag, will tend to fall (joint supply). More munitions factories will have to be built (joint demand), and this will create a shortage of bricklayers, so upsetting the housing programmes of local authorities (composite demand), and forcing them to build timber and concrete houses (composite supply). At the same time rearmament will give a fillip to all the industries supplying raw materials for arms, from the producers of optical lenses for service binoculars to the makers of buttons for uniforms (derived demand). The reader will find it instructive to work out less obvious repercussions; to trace, for example, how recruiting and aerodrome construction create a shortage

of agricultural labour, increase the demand for tractors and milking machines, and reduce the demand for horses and oats.

When two commodities are interdependent, on the side either of supply or of demand, the task of disentangling the cost and utility of one from the cost and utility of the other might appear to be almost impossible. We cannot produce wheat without straw; what, then, is the cost of the wheat and what the cost of the straw? If we know the joint cost of wheat and straw taken together, how are we to assign this cost between the two? This question can be answered satisfactorily only if we understand why it is ever asked. Information about the cost and utility of a commodity is of value only if it helps us to decide whether too much or too little of the commodity is being produced. That is, the information ought to bear on *marginal* units of the commodity and show whether there is a balance of gain or loss in producing these units. It is vital, therefore, for a firm to know the *marginal* cost of each of its products. But to divide up joint costs and distribute the total between each joint product is quite meaningless. The allocation of joint costs, while it may decorate a balance sheet and flatter the accountant's sense of propriety, serves no purpose whatever.

How, then, can we calculate the marginal cost of wheat? If wheat and straw are always yielded in a fixed ratio, so that we cannot increase the supply of one *at the expense of the other*, the answer is quite simple. The marginal cost of wheat is what can be saved by not growing a little of it, and this necessarily includes the cost of that inevitable concomitant of wheat-growing—straw. Similarly, the marginal cost of the straw includes the cost of the wheat that grows on it. The marginal cost of each is equal to the joint marginal cost of both. At the same time the marginal revenue of wheat will include the value of the straw which is produced simultaneously, while the marginal revenue of straw will include the value of the wheat which is produced simultaneously.

This may seem a very complicated way of explaining that the cost of wheat is neither separable nor different from the cost of straw. But observe how illuminating the explanation becomes if we *can* vary the proportion in which wheat and straw are grown—if, for example, we can grow wheat on a longer stalk or with a heavier ear. A marginal addition to the supply of wheat can then be made without sowing a larger acreage, for example, by sowing varieties of seed, or using types of fertilisers, or taking the kind of trouble over cultivation that will yield more wheat on a given weight of straw. There will be various marginal adjustments, that is, each of which will increase the yield of wheat without increasing the yield of straw proportionately—adjustments which, in other words, will substitute wheat for straw. Suppose that these adjustments are made in such a way as to yield the same quantity of straw as before and an additional bushel of wheat. Let us select that particular device by which this can be contrived at lowest cost. Then this lowest cost of producing an extra bushel of wheat gives us the marginal cost of wheat. When

the marginal cost of wheat can be calculated in this way it is obviously quite distinct from the marginal cost of straw.

This is not a full solution to our problem. The price of wheat, even in the long run, may remain below marginal cost as calculated above. For although it may be *possible* to grow an additional bushel of wheat on the same quantity of straw, this may not be the most *profitable* expedient. It may be easier to grow a little more straw, and a good deal more wheat, setting the value of the additional straw against the increase in total costs. The price of wheat will then be governed by the rule that marginal cost tends to equal marginal revenue, the first including the cost of any additional straw and the second its value.

**Joint Supply and Composite Demand.**

We have seen that since wheat and straw are joint products, an increase in the demand for one will tend to bring down the price of the other. But this conclusion was reached on the assumption that more wheat could not be produced without the simultaneous production of more straw. If the proportion between wheat and straw could be easily varied, the conclusion would be reversed, for it would be possible to grow more wheat by growing *less* straw and so making it dearer instead of cheaper. There would, in fact, be a *composite demand* for the wheat crop, and wheat and straw would be alternative rather than joint products. When two commodities are produced together, therefore, there are two aspects to the association. When there is great difficulty in varying the proportion in which they are produced, we say that they are in joint supply; when the proportion is easily varied, we say that there is a composite demand for the factors engaged in their production. If the demand for one commodity increases, the price of the other will rise or fall according as the aspect of composite demand or joint supply predominates. To put the same conclusion rather differently: the greater the ease with which the production of A can be abandoned in favour of the production of B—in technical terms, the greater the elasticity of substitution of A for B—the stronger will be the tendency for prices of A and B to move in the same direction.

**Joint Demand and Composite Supply.**

On the side of demand symmetrical conclusions apply. Goods in joint demand are bought in fairly constant proportions, while goods in composite supply, being substitutes for one another, are bought in highly variable proportions. If the aspect of joint demand predominates, an increase in the supply of A tends to raise the price of B; if there is more whisky to be had, we shall want more soda and probably have to pay more for it. On the other hand, if the aspect of composite supply predominates, an increase in the supply of A tends to reduce the price of B; if tea is more abundant, we shall be less anxious to buy coffee, and its price will fall. The greater the ease with which we can substitute one commodity or another—the

greater the elasticity of substitution by consumers—the stronger will be the tendency for the prices of the two commodities to move in the same direction.

**Joint Products from the point of view of the Firm.**

The practical implications of these ideas may be clearer if we apply them to the problems that face the individual firm. Suppose, for example, that some machine, or some generator of power, or some piece of floor space is contributing to the output of several commodities. Then it will be a waste of time to try to allocate to each commodity its share in the joint or "overhead" costs[1] of the machine, or the power plant, or the floor space. But it will be of the utmost importance to find what saving could be made on these overhead costs by reducing the output of any one commodity, and what additional outlay on overhead costs would be necessary in order to increase the output of that commodity. If there is no *measurable* saving or additional outlay attributable to individual commodities—if changes in overhead costs affect all commodities equally—then we must try to find out, on the assumption that overhead costs remain unchanged, what reduction in the output and sale of *other* goods is necessary in order to provide for an expansion in the output and sale of each individual commodity. For example, a shop which sells butter and eggs, and has a fixed amount of floor space, may be able to estimate the loss in revenue from butter that will result from the stocking of more eggs and less butter. This loss in revenue can be regarded as the marginal cost of retailing eggs and should be weighed against the gain in revenue from the additional eggs that are sold. The problem of charging butter and eggs with their appropriate share of the cost of floor space (the rent of the shop) simply does not arise.

Butter and eggs are not only in joint supply, but also in joint demand—that is, many customers normally purchase both in the same shop. A grocer who reduces the price of eggs, therefore, will not only increase his sales of eggs, but will also attract custom in other lines, such as butter, and the profit which he makes on his additional sales of butter will go to reduce the margin required to cover the cost of retailing eggs. If eggs are sold at a specially low price, the sacrifice of profit on eggs may really be a species of advertisement designed to increase profits from the sale of other groceries.

This parallel with advertisement and selling costs suggests that the principles governing selling costs apply equally to joint products. As before, once we define marginal cost and marginal revenue broadly, the rule that they tend to equality with one another holds. The position of maximum profit, when several products are being produced or several markets supplied, will be one in which there is

[1] Notice that in this context overhead costs are "costs common to a number of commodities," not "costs fixed over a period of time."

no commodity such that by producing more of it the additional net revenue exceeds the additional net cost. The net (marginal) revenue will include the rise in the proceeds of sale of the commodity, less the fall in returns from any commodity the output of which has had to be cut down, plus the rise in returns from any commodity which is in joint demand with the first, etc. The net marginal cost will include net additional outlay plus any additional depreciation due to the expansion of output.

The same line of reasoning applies to the discontinuance or initiation of the production of some commodity. If there is a prospect of additional profit, taking into account all extra outlays and the balance of increase in receipts, then the commodity is *ipso facto* profitable. The problem of saddling the commodity with its "share" of overhead costs, which are perhaps unaffected by this new departure, does not arise.

When the question of renewal presents itself, the considerations on which a decision should be based are similar. A machine may be used in the manufacture of several commodities. The manufacturer will, therefore, have to sum the various reductions in his receipts which failure to renew will involve and weigh the total against the annual charge for upkeep, depreciation and interest on a new machine. If there is some other method of production which promises larger profits, or lower losses, then the machine will not be renewed.

The fundamental point is that costs must be approached from the point of view, first, of alternatives, and, secondly, of the margin. The changes which have to be analysed are generally marginal changes—repercussions of a suggested change must be worked out and the result compared with the existing position. The profitability of any output is thus purely relative. It may pay to make an apparent loss because an offsetting gain is being made, or is expected to be made, in consequence of the loss. To insist on making a profit everywhere may result in reducing profits all round.

**Price Discrimination.**

The practice of dumping, or price discrimination, offers an excellent illustration of this truth. Dumping may arise whenever the products of a firm are in composite demand by distinct groups of consumers, each group being made to pay a different price for exactly the same product or service. The Milk Marketing Board, for example, charges one price to consumers of liquid milk, and a range of much lower prices for milk to be used for manufacture into butter, cheese, milk powder, and other milk products. Now it can readily be shown that if a firm which is in a position to discriminate insists on making a profit in each market, its total profits may be less than they would be if the firm were satisfied with an apparent "loss" in some of its markets. Suppose, for example, that the firm possesses semi-monopolistic powers in one market (e.g., the home market) and is open to severe competition in another (e.g., export markets) so that it is more reluctant to spoil the first market than the second. The

price in the first market may be £40, and in the second £18, while sales are 200 in the first and 300 in the second.

| | Price | Sales | Total Receipts | Average Cost |
|---|---|---|---|---|
| First Market | £40 | 200 | £8,000 | £30 (200 units) |
| Second Market | £18 | 300 | £5,400 | £20 (500 units) |

Suppose also that the average cost of production of 200 units is £30, while the average cost of (200+300) units is £20. Then it might appear that every unit sold in the second market would lose the firm £(20 — 18) or £2. In fact, however, refusal to sell in the second market results in the loss of £5,400 in revenue, and saves only £(500×20) — (200×30) or £4,000 in cost. In other words, selling below (average) cost in the second market increases total profits by £1,400.

Once again, thinking in terms of marginal cost and revenue saves us from confused reasoning. The marginal cost of output will depend mainly on the total amount produced and will be more or less independent of variations in the proportion disposed of in any one market. Now it will pay the manufacturer to make marginal cost and marginal revenue equal in *each* market, and since the gap between marginal revenue and price will be least in the highly competitive export market, he will tend to charge a lower price there than in the home market. If demand in the export market is so elastic that his power to depress or raise prices is negligible, then he will expand his sales abroad until marginal cost is equal to the world price, and since marginal cost will remain below average cost so long as costs are falling, price will also be lower than average cost.

It should be observed that the incentive to discriminate originates not in unused capacity as is sometimes suggested, but in the imperfection of competition. If competition acted with equal force in all markets, manufacturers would charge the same price all round, whether their costs were rising or falling. If it were impossible to separate one market from another, the same would be true. Discrimination can only be resorted to whenever groups of consumers can be separated out by geographical area, or income, or membership of some society, or propensity to visit particular shopping centres, or to buy particular qualities of product. There must also be some barrier to movement from one group to another. For example, goods dumped abroad may, if the price disparity is excessive, be reimported and bring down prices at home. In the same way, if shopkeepers charged red-haired customers twice as much as dark-haired customers, the victims—unless red hair became a sign of superior social status—would dye their hair and escape victimisation.

**Joint Products from the point of view of Society.**

From the social point of view, all products are, in the broad sense, in joint supply ; the stream of goods and services that makes up the national output is the joint product of all our resources of land, labour and capital. From the social point of view, all products are also in joint demand ; the stream of income that makes up the national income provides, when expended, the joint demand for the products and services of the factors of production. The national income is, in fact, simply the national output upside down, or looked at from the point of view of the consumer instead of the producer. What we produce flows into a reservoir in which are pooled the products of our neighbours. What we consume is drawn from the same reservoir, from the joint output of the community.

Thus whereas the individual firm, in balancing cost against price, can confine itself to the limited range of products which it manufactures ; and whereas the individual consumer, in balancing price against utility, can confine himself to the limited range of products which he buys, calculations of *social* welfare must embrace *all* goods and services. In private accounting there are many repercussions of our decisions to buy or sell that we can, if we choose, neglect ; but in social accounting, since all products are joint products, none of these repercussions—unless they are trifling—can be neglected. The producer of chemicals, whose waste gases pollute the atmosphere, does not reckon in pollution as part of his output. The consumer who burns raw coal in the domestic grate does not deduct the soot and dust which he creates from the utility of the fire. Private costs and private preferences are partial and incomplete. Our private interests and the interests of society do not coincide.

To these points we shall return in the next chapter when we consider whether prices in a capitalist society adequately reflect the wants of the community.

## CHAPTER 18

## SOCIAL ASPECTS OF PRICING

NOW that we have seen how prices are fixed, it is time to turn to the far more perplexing problem how prices should be fixed. A useful point of departure is provided by the suggestion which has been made by some economists that the ideal working rule is to charge prices equal to marginal cost of production. This rule, and this alone, it is argued preserves the right balance between value and cost. The price of a commodity measures the value which consumers set on an extra unit of it ; marginal cost measures the cost of producing such a unit. If price is above marginal cost, therefore, too little of the commodity is being produced ; if price is below marginal cost, too much of it is being produced. The output of each commodity will be ideal only if price and marginal cost are everywhere equal.

There is one obvious objection to this rule: in most industries it could be adopted only at a loss. A profit can be made only if average cost is *below* price. But average cost is generally *above* marginal cost, to which it is proposed that price should be equated. The rule, therefore, might appear to be unworkable, since it would involve most industries in loss. If only one or two industries were affected, the difficulty might be overcome by a system of subsidies and taxes calculated to equalise profits. Those industries in which marginal cost—and therefore price—was above average cost would be forced to disgorge their profits in taxes, the proceeds of which would enable other industries to cover their costs. This device, however, is insufficient when more industries are making losses than are making profits. In such circumstances we must relax the rule to make price, not equal to, but proportional to, marginal cost—charging prices in each industry in a fixed ratio to marginal cost so as to bring in, on the average over the whole of industry, whatever profit is considered to be adequate. Some industries will still be involved in loss, and will be granted subsidies in order to induce them to continue production of the ideal output, while other industries will have their excess profits skimmed off in taxation.

It might be supposed that this system of levies and subsidies, instead of making the output of each industry ideal would make it indeterminate —for each industry seems to be assured of a normal profit *whatever* output it produces. An industry, one might imagine, could expand output with impunity, feeling itself guaranteed against loss, and varying its price to match changes in marginal cost. But this is not so. The industry must *sell* the whole of its output, and there is only one point on the demand curve—and only one output, therefore—which satisfies the pre-determined ratio between price and marginal cost. If, for example, output is doubled and marginal cost falls by 10 per cent. there is no reason to expect that sales will be doubled by a 10 per cent. reduction in price.

The real difficulty is an administrative one. The rule could only be adopted if very wide powers were entrusted to the State—far wider powers than could easily be grafted on to capitalist institutions. The State would require to have accurate knowledge of marginal costs in each industry—knowledge which we do not at present possess, which it might be extremely difficult to obtain and which, if available, might be open to conflicting interpretations. There would require to be frequent adjustments in rates of levy and subsidy—deliberate adjustments which would meet with vigorous resistance from the industries affected. It is one thing to leave the conflicting interests of each industry to the impersonal arbitrament of the market, and cling to the maxim of capitalism, that industries which cannot make a profit do not deserve to survive. It is quite another for the State to take the responsibility of deciding that this industry should expand or that contract. A fall in the price of coal can be represented as something beyond human control, which the mining industry must accept in a spirit of fatalism. But a government decision

that too much coal is being produced is seen in a quite different light.

These difficulties are clearly not confined to capitalism—a Socialist State would be open to pressure of the same kind. But it would not be in the position of trying to run industry on two conflicting principles simultaneously—that of maximising profits and that of producing an ideal output. So long as, over the whole field of industry, costs (including the costs of uncertainty[1]) were covered, a Socialist government would be free to charge such prices as it thought fit, and it could therefore adopt the rule of making prices proportional to marginal costs with as much consistency as political pressure allowed.

The main consequence of applying the rule, in existing circumstances, would be to cause a transfer of resources from industries in which there was a low " degree of monopoly " to industries in which the " degree of monopoly " was comparatively high. In those industries in which competition approached perfection—and in which, therefore, price approached marginal cost—a comparatively large rise in price would be necessary in order to bring the ratio between price and marginal cost into line with other industries. At the same time, industries with great powers of monopoly would require to bring down the price of their products. The first group of industries, faced with a contracting market, would require to restrict output and release resources ; the second group, finding demand increasing, would expand output and employ additional resources. Monopolised goods, previously too scarce, would become more plentiful ; and the resources necessary for their production would come from industries in which output had previously been too great. The transfer of resources would be not only to monopolies as generally understood—monopolies based upon a restriction of entry. There would be a transfer also to the manufacture of goods protected from competition by narrowness of the market—that is, to goods manufactured on less than their optimum scale of output. On the one hand, monopolistic restrictions would be removed and all industries put on the same competitive footing ; on the other hand, fuller advantage would be taken of the economies of large-scale production. We should have more and cheaper electric lamps, cotton thread, milk and similar monopolised goods ; and—far more important—we should have more and cheaper refrigerators, motor cars, railway journeys, electricity, and all the other things most subject to economies of scale.

A second consequence of applying the rule would be to put an end to price discrimination. Each consumer would pay at the same flat rate for any standard commodity or service. There are, however, some circumstances in which departures from the rule might be made either in the interests of poor consumers or with a view to using up a surplus. There would be no reason, for example, to give up the existing practice of providing milk free or on special terms to school-

[1] See page 273.

children, unemployed workers, expectant mothers, and similar groups of consumers. And it might be arranged to distribute surplus fruit, fish and other foodstuffs at low prices to consumers who were generally unable to purchase enough of them and who could be induced to consume more without upsetting the market. Similarly, isolated villages might be allowed to use the telephone at less than cost, and electricity might be installed in working-class houses at specially low rates. Discrimination, in short, might be used deliberately as an instrument of social policy so as to direct the demand of groups of consumers towards certain essential commodities or towards commodities especially subject to economies of scale.

The rule that price should be proportional to marginal cost is intended to secure that the wants of consumers receive the fullest possible satisfaction—if we are willing to sacrifice a little of A for the sake of a little more of B produced at equal cost, then our preference will be met by an appropriate transfer of resources from the production of A to the production of B. But what if we want to sacrifice a little of A for the sake of something not on the market at all? We cannot give effect, as individual consumers, to a preference to a less varied and cheaper selection of goods; or for economic progress rather than economic security; or for an increased National Debt rather than increased unemployment. We have freedom of choice, but only within the limits of existing market conditions. If we want to change market conditions we must associate with other like-minded consumers in bringing pressure on the Government. Our preferences find expression, not in the market, but in Parliament. But they are none the less preferences of which account must be taken in deciding what is to be produced and how it is to be produced. Preferences backed by voting power can be just as important in arriving at a just price or an ideal output as preferences backed by purchasing power.

Nor is this a trifling difficulty. For the deepest wants of man have little to do with the scarcity which economic activity is primarily designed to combat. Man craves most of all a sense of purpose, a rôle in life and in society. Deny him that, make him feel of no consequence, and he will take revenge by overturning any economic system, however well regulated. It is not enough to offer him value for his money in goods and services. He must have a place to match his sense of merit—something to admire and defend, entertainment and amusement, responsibility and initiative. Without these, goods and services are as dust and ashes, and any one output is as far from ideal as any other.

So long as we refuse to go behind man's wants to the inner motives which inspire them, the rules of pricing can be reduced to simple axioms. We can treat each want atomically as a demand for this or that service, and discuss choice as the balancing of one want against another. But immediately we begin to think of wants as fused by a sense of purpose, the tool of marginal analysis breaks in our hands. It is borne in on us that it is not just the surface wants that cry out for satisfaction, but the deeper urges which they often express. The

African chief who, in fear of death at the hands of his tribe, seeks hair-dye of white travellers to prolong his youth and his rule[1]: the *nouveau riche* who buys a country seat so as to impress his associates: the purchaser of cosmetics, parrots, or holy relics: all have resort to market agencies in order to satisfy wants that are denied satisfaction in other ways. They buy goods not for their own sake so much as for what they symbolise. And this is true not just of a few odd people or of a few odd purchases, but of practically everyone at some time and of the bulk of the purchases which most of us make at any time. We do not buy the cheapest article available and only those articles which are indispensable to living. We buy what appeals to our imagination under the pressure of social standards. In economic activity as in other spheres of life the motives which dominate us are rarely hunger and want in the narrow sense (except at a very low standard of living) but the desire for power, the desire to make others feel that power, to show off to and keep up with, our neighbours, to plan and build and dream. We seek distinction, for example, in our dress or in our profession; and power over others through our wealth or as their employers. To isolate individual buyers and sellers in some vacuum of choice, therefore, and discuss how industry can best maximise an ethereal utility is to misconceive the problem. We might as well discuss what combination of medicines will most promote health, without regard to the causes of ill-health or to the other remedies (fresh air, exercise, good food, adequate clothing, and so on) that lie to hand. To substitute one assortment of goods and services for another may do far less for the people's welfare than the setting of new social standards or the direction of their energies into new and more acceptable channels.

Nor is this all. For the act of substitution of itself *changes* social standards and *creates* new demands. If, for example, we provide cheaper wireless sets we may reduce the demand for theatre seats and increase the demand for coal, since more people will wish to stay at home and listen in. Thus the balance between value and cost is upset and no one can say with assurance whether it is upset for better or for worse. We cannot assume, therefore, that in fixing outputs so that prices are in a fixed ratio to marginal cost we are giving fuller satisfaction to the wants of consumers, for these wants are themselves affected by variations in output.

Again, many of our wants are shaped by the very system of production which exists to supply them. A man's habits of living and of expenditure depend partly upon the work which he is called upon to do. A worker in casual employment has a different scale of values from a worker in regular employment. A factory worker, a miner, an agricultural labourer, a stockbroker, and a doctor have each their own outlook and make their own peculiar purchases. By changing methods of production, therefore, we can simultaneously change the demand of consumers. If we provide well ventilated and

[1] Aldous Huxley: "Texts and Pretexts," page 142.

clean factories we can create a demand for well ventilated and clean houses. If we cut down overtime we may be able to reduce the demand for medical services—or for whisky. Or if we take the economic system as a whole, one kind of economic organisation may create a greater danger of war than another because of the habits of mind which it inculcates. The cost of war is far from negligible; but the cost of producing the mentality or the social arrangements which would get rid of war does not figure amongst the costs which it is proposed to equate to price.

In short, we cannot base prices solely on current market valuations because the money costs and values which find expression in the market are incomplete measures of social costs and values. Buying and selling go on against a social background from which we cannot abstract in estimating the value of what is bought or the cost of what is sold. Consider, for example, the building of a house. It costs, let us say, £1,000 and is valued at the same figure. But do we enter up in the cost the noise which the neighbours endure, the nuisance to passers-by from added traffic congestion, or the loss of amenities to other householders? Do we include in the value of the house its effect on our health and on the health of our children (matters of immediate interest to the community, which may have to support us if we fall ill, has to combat any infectious disease which we may help to spread, and is out of pocket on the cost of our education if we are unable to do regular work and to pay our taxes)? Do we include the effect which living in such a house has on our character and so on our work and friendships? Do we include the pleasure which other people derive from our garden? It is highly unlikely that we include any of these things. Yet they affect the cost and value of the house to the community just as surely as the cost in bricks and mortar and the number and size of the rooms.

Or let us take an industry—agriculture or coal-mining. Is the contribution of a farm or mine to human welfare exhausted in the food or the coal which it produces? Surely not. The stability of agricultural employment, the training in skill and responsibility which farming provides and—not least—the amenities which it creates cannot be omitted from the balance sheet of the industry as the minutiæ of value. It is only when all these advantages have been reckoned in, and a broad judgment of social welfare superimposed upon our working rule, that we can decide on the ideal output of agriculture—or of any other industry. Market forces give expression only to marginal utility and marginal cost as judged by buyer and seller, not to social utility or social cost. Thus the machinery of the market—even when regulated so as to keep marginal utility and marginal cost equal to one another—is insufficient to make the output of each commodity ideal. We must supplement market competition, with its narrow perspective of self interest, by some other kind of competition from which a broad judgment of social welfare will emerge. In democratic countries, this second kind of competition takes the form of competition for votes. It is the electorate which,

in the last resort, has power to decide whether one industry is too small or another too large to be in the social interest.

A wider question remains. Is it right to take demand for granted and give it the fullest possible satisfaction, without first deciding whether or not it *ought* to be satisfied? Clearly, there is no ideal output of drugs. But what of hundreds of other commodities which consumers buy ill-advisedly or ignorantly? Is the consumer the best judge of his own interests when he buys patent medicines, or petrol, or electric fittings? Still more important, should the value which he sets on milk or on opera or on education be the sole criterion by which the output of these things should be determined? In general, it is wise to leave the consumer to learn from his own mistakes and not to dragoon him too insistently into paths of wisdom and virtue prepared by a very fallible government. But there are times when, in fixing prices, the State may properly disregard individual valuations as shortsighted, mistaken or wrong, deny access to certain commodities in the quantities which consumers would be willing to absorb, and offer other commodities in larger quantities, in more attractive surroundings or backed by more advertising pressure, than free competition would provide. If technical experts are practically unanimous in praising one commodity and in condemning another, the State will be justified in setting aside any contrary verdict by individual consumers. It will be right to encourage, or even to enforce consumption of the first, and to discourage or suppress the sale of the second. When informed opinion is seriously divided, however, or when important moral issues are involved, the State will be unwise to interfere unnecessarily with the free choice of consumers in order to impose upon them ideals which they do not share. To prohibit the drinking of methylated spirits is one thing; to prohibit the drinking of beer, quite another.

**Conclusion.**

The task of pricing is part of a wider task of social organisation from which it cannot, and should not, be divorced. No rigid formula can resolve the subtle and complex problems which it raises. Nevertheless, as a working rule, there is much to be said in favour of keeping prices in a fixed ratio to marginal cost, provided that deviations are permitted when broad social interests so require.

# PART IV—THE DISTRIBUTION OF INCOME

## CHAPTER 19

## THE DISTRIBUTION OF INCOME

THE theory of distribution might be expected to deal with the reasons why some people are rich while others are poor, and whether there is any justification for these inequalities. Traditionally, however, economists have preferred to put the question in a different way. They have analysed the distribution of the national income, not between persons, but between the factors of production—land, labour and capital. They have set themselves to explain how rent, interest, profits and wages are fixed, leaving over for later study the resulting distribution of income between the owners of the factors of production. This procedure splits the problem of distribution into two parts. The first, which is an extension of the theory of value, analyses the forces governing the *prices* of the factors of production; it is a study of the tendencies of competition. The second, which involves research into the social structure of particular communities rather than the working out of economic principles, analyses the forces governing the *ownership* of the factors of production; it is mainly a study of the influence of inheritance and political power. Once we know what governs both the earnings and ownership of the factors of production, we know also what governs the distribution of income between persons and we can go on to discuss how far this inequality is either necessary or justifiable.

In this chapter we shall concentrate on the first problem. We shall assume, to begin with, conditions of perfect competition. We shall also assume conditions of full employment—meaning by this that the economic system is functioning so as to allow each factor as much employment as its owner wishes at the current rate of remuneration. By making these assumptions, and setting aside the influence of monopoly and unemployment, we are able to concentrate on a single issue: What are the tendencies of competition?

The price of any factor of production, like the price of a commodity, is governed by supply and demand. The scarcer the factor, and the greater the demand for its services, the higher will be its earnings. The more abundant the factor, and the less urgently it is required, the lower will be its earnings. But this does not take us very far. What is it that governs supply and demand?

**Supply.**

In early theories of value no importance was attached to demand. The only lasting influence on price was that of supply—in other words, of cost of production. This was true not only of commodities

but also of the factors of production; and just as an exception was made in favour of "rare statues and pictures, scarce books and coins," where competition was powerless to increase the supply, so also an exception was made in favour of land, where again the supply could not be increased. The price of old masters depended solely on the demand; so also did the rent of land. But interest and profit, and above all wages, were governed by supply. The Subsistence Theory of Wages, for example—the so-called "Iron Law of Wages"—was simply a statement that the price of labour tended to be just sufficient to cover the cost of production of the labourer. Wages, it was thought, must oscillate round a level that was just high enough to afford the wage-earner a bare subsistence. Higher earnings would induce workers to have larger families, the labour market would become overstocked, and wages would be forced down again. Similarly, if wages fell below the level of subsistence, workers would die off and the birth-rate would fall until the ensuing shortage of labour raised wages again.

This gloomy theory rested largely on the work of Malthus. In his "Essay on Population," published in 1798, Malthus argued that the human race had the power—and the propensity—to multiply its numbers more rapidly than the means of subsistence, that population was in fact increasing, and that this increase would inevitably lead to pressure on the means of subsistence resulting in a high death-rate through famine, war and disease, unless mankind curbed its unruly instincts in later and less fertile marriages. The middle classes, convinced that the instincts of the poor could never be curbed, found Malthus's doctrines singularly soothing. If money given to the poor only encouraged them to multiply, then the money was far better in the pockets of the would-be philanthropists—and their numerous children. Those who, sharing Malthus's views, struggled to raise the standard of living of the workers, must have felt that they were building on sand and that no real advance could be made so long as the population kept on increasing. It is significant that the development of the public social services in Britain made little progress until the views of Malthus lost favour.

The experience of the nineteenth century proved that Malthus's fears, although not altogether groundless, were exaggerated. The population of England and Wales, it is true, increased fourfold. But instead of an accompanying fall in wages below the level of subsistence, wages too increased fourfold; and instead of a rise in the birth-rate to match the increasing prosperity of the working class, the birth-rate began to fall and kept on falling. Far from high wages increasing the "supply" of wage-earners, the reverse appeared to be true; the birth-rate was lowest in countries where wages were highest. Thus the Malthusian theory was turned upside down. Measures to raise the standard of living of the poorer classes would relieve, not aggravate, the pressure of population.

The fact is that, although the growth of population does depend upon the level of wages and incomes, the connection is by no means

a straightforward one, and there are many other influences quite unconnected with wages. We may study these influences and try to forecast the future of population, and we may study also how changes in population will react on wages. But this is a very different thing from laying it down that whatever happens to population, wages will tend to sink to the level of subsistence and that the demand for labour can, therefore, be neglected. A cost theory of wages along these lines must be ruled out.

**Real Costs : Efforts and Sacrifices.**

Wages are not related very closely—at any rate in an industrialised community—to the cost of "producing" labourers (i.e., of rearing children). But what of the cost of inducing labourers to work, putting up with hardship and discomfort for long hours with short holidays? Can we not say that each hour of work costs the worker something in the effort which he expends and the sacrifices which he undergoes? For those efforts and sacrifices wages are presumably intended to compensate, and high wages, therefore—whatever their effects on population—may be expected to call forth a greater supply of work by disposing the worker to make greater efforts and greater sacrifices. We decide how much work we are willing to do by weighing the wage offered against the subjective "cost" of working; the wage must be high enough to overcome our disinclination to undertake disagreeable tasks. Indeed, we can go further. Provided we are free to choose which tasks we will do, the wage must be high enough to compensate us for the *most disagreeable* tasks. We balance the additional earnings which we can make against the additional hardship of working for one more hour, or at higher speed, or in less pleasant surroundings. It is the *marginal* disutility of work—the disagreeableness of those tasks which we are most tempted to decline—that governs the supply of labour.

This view of costs has two important disadvantages. In the first place, it suggests that a rise in wages will call forth greater effort from wage earners. But it often happens that a rise in wages, by making it easier for workers to reach their customary standard of living, induces them to work less energetically, or take longer holidays, or absent themselves more frequently from work. It is well known, also, that in times of trade depression when wages are falling, workers often respond by increasing their output, partly for fear of losing their job, but partly also in an effort to maintain the old level of earnings. The more firmly workers cling to some fixed standard of living, the more likely it is that the supply of labour will move in the opposite direction to wages.

Secondly, the linking of wages with real costs suggests that wages are somehow *proportioned* to real costs. But it would be ludicrous to suggest that the duties of a professor are ten times as disagreeable as those of a coal-miner or blastfurnaceman, or that his income is ten times greater than theirs because of the greater efforts and sacrifices which his training required. Earnings may be related, but are

certainly not proportioned, to efforts and sacrifices. This is particularly obvious if we look at the earnings of the other two factors of production. The efforts and sacrifices for which labour is compensated are real enough. But in return for what efforts or sacrifices do we pay interest on capital or rent for land? The provision of land, as we have seen, involves no sacrifice because the land already exists and cannot go out of existence. The provision of capital does, however, involve a sacrifice since we can add to the stock only by saving, and saving means denying ourselves present enjoyments. It costs us something to refrain from spending money, and the interest which we earn on our savings can be regarded, therefore, as compensation for our self-denial. But no one would dream of suggesting that this self-denial is as great for the millionaire as it is for the wage-earner, although both are paid equally at the same rate of interest. In addition, the stock of capital that was accumulated by past generations requires no fresh sacrifice, although it continues to earn interest. Clearly, then, the payments which are made to the factors of production bear little relation to the sacrifices by which they are earned. Rent is earned without effort on the part of anyone, interest by dint of thrift or in virtue of inheritance, and wages by the sweat of the brow. The service sold by one man for a pound may cost him a far lighter sacrifice than the service sold by another man for a shilling.

**Supply Price: Costs as Relinquished Alternatives.**

Payments to the factors of production are made, not in compensation for sacrifices, but as inducements to effort. These payments—the prices of the factors—must be high enough to provide an adequate incentive to the owners of the factors to continue to supply them. Each factor must earn its "supply price"—the minimum price, expectation of which will just suffice to call forth the required amount of the factor—or its services will not be made available. This supply price will depend upon the pull of the alternatives open to the factor. If an hour of leisure is very attractive in comparison with an hour of work, then the supply price of labour will be correspondingly high. If we are little troubled by thoughts of a rainy day in the distant future, and weak against the fascinations of present enjoyments, then it will require a very high rate of interest to turn the scale in favour of thrift. Our time can be spent in work or in leisure, and our money can be spent or saved. The greater our preference for work, the lower will be the supply price of labour. The greater our preference for thrift, the lower will be the supply price of capital. The strength of our preference will make itself felt at the margin. The choice before us is one of a little more work or a little less, a little more spending or a little less. Thus it will be our marginal preference for work, as compared with leisure, which governs the supply of labour (in the sense of actual effort), and it will be our marginal preference for thrift which governs the supply of capital (in the sense of current savings).

The rate of wages and the rate of interest must be at least equal to the supply prices of labour and of capital, and these supply prices are governed by our marginal preference for work and for thrift. In other words, the rates of wages and of interest measure the attractiveness at the margin of the alternatives which we relinquish by working or by saving. The value of these alternatives is the true cost of working and saving.

**Supply Price and Transfer Cost.**

This conception of cost as relinquished alternatives is one with which we have already met. We saw how the cost of producing any commodity is equal to the value *for other purposes* of the factors engaged in producing it—what we called their transfer cost. Now "transfer cost" is just another name for "supply price." But whereas in the present chapter we have been discussing the supply price of labour and capital *for any purpose*, what we discussed previously was the supply price of labour and capital to some *particular industry*. There is a distinction between the question: What determines the total amount of labour and capital seeking employment? and the question: What determines how a given amount of labour and capital will be distributed between competing industries?

This distinction is of great importance when we come to consider the third factor of production—land. The supply price of land-in-general is zero. There is no alternative to using land except not using it, and since there is nothing to be gained by not using it, its supply price is zero. The entire earnings of land, therefore, form a surplus above its supply-price and it is by analogy with land that any excess of earnings over transfer cost or supply price is called "rent." On the other hand, the supply price of land to any one branch of agriculture is by no means zero. If, for example, land can be used either for wheat growing or for barley growing, then the supply price of the land to either industry will be its value under the other crop. The cost of using the land for wheat growing, therefore, is quite different from the simple cost of using the land; the alternative to be overcome is the comparatively attractive one of using the land for barley growing, not the quite unattractive one of letting the land go out of cultivation.

**Rent and Cost: An Illustration.**

To illustrate this point in more detail it is necessary to elaborate the theory of economic rent. Economic rent is the payment made by a tenant for the use of land alone. The rent paid by a British farmer (contract rent) is generally greater than economic rent since it includes a payment for the use of buildings and fences and for other improvements such as clearing and ditching. Economic rent also differs from contract rent in that it continues to exist whether the owner of the land lets it to someone else or works it himself.

The essence of the theory of economic rent as it used to be stated—for example, by Ricardo—is that rent "does not and cannot enter in

the least degree "[1] into price, but is itself governed by price. Clearly, this is true only if we are thinking of land-in-general. If it were possible to use land for a single purpose only, say for wheat growing, then the rent of land would have no influence upon the price of wheat. The supply price of land would be zero—that is, there would be no level of rent below which land would cease to be available for wheat growing. Rent, therefore, would be a surplus governed by the demand for wheat and the cost of cultivation. It would be high if wheat was fetching a high price, or if the cost of cultivation was low, so that there was competition for farms; and it would be low if an agricultural depression, by bringing down the price of wheat relatively to farming costs, made it difficult for farmers to carry on. If the price of wheat were high it would not be because high rents had to be paid. The reverse would be true; it would be because the price of wheat was high that high rents could be paid.

This proposition is most readily understood when it is related to the idea of the margin. The price of wheat will be equal, under perfect competition, to the marginal cost of growing wheat—in other words, to the cost of growing wheat on marginal land. If the price is higher, additional land will be brought under cultivation, and if it is lower, some land will fall out of cultivation. But marginal land pays no rent; and rent, therefore, cannot form part of the cost of cultivation at the margin. It is this cost which tends to equal the price of wheat; rent, therefore, can have no influence on the price of wheat. Rent will be paid by intra-marginal pieces of land, on which the cost of cultivation is lower than at the margin, and since farmers are free to choose between marginal land and land of greater fertility or superior situation, the rent which they will offer for any piece of land will measure the differential advantage which it possesses over land at the margin. If the demand for wheat increases, forcing up the price and making it profitable to cultivate, at increasing cost, land which was previously sub-marginal, the differential advantage of a given piece of land over marginal land will automatically increase and the rent which is offered for it will increase correspondingly. But the rise in rent will be the consequence not the cause of the rise in wheat prices. The price of wheat would not be affected if farmers paid no rent whatever.

The same conclusion holds of urban rents. Goods in Bond Street are dear not because the shops there pay high rents, but because customers are willing to pay high prices. These high prices increase the demand for shop sites in Bond Street, and so enable rents to be paid. Bond Street rents, in other words, reflect the differential advantages of sites in the centre of London over sites on the outskirts.

Suppose, however, that there are other crops than wheat and that Bond Street shops do not sell goods of the same kind. Then the supply price of land *for any one purpose* ceases to be zero. Land can

[1] Ricardo: "Principles of Political Economy" (Everyman Edition), pages 40-41.

be rented for wheat growing only if its use is denied to other crops, and the rent which it might yield under the most profitable of these other crops—the transfer cost of land—does enter into the cost and into the price of wheat. If this cost is not met, the land will cease to be available for wheat growing and will cross the margin of transference into the next most profitable use. Land on the margin of cultivation pays no rent; land on the margin of transference does pay rent. This rent enters into the cost of *particular* agricultural products; it is not a surplus over the cost of cultivation, but is itself part of the cost of cultivation, governing, not governed by, price.

The general level of rents, therefore, is governed by the demand for agricultural products as a whole and by the cost of cultivation at the margin; from the point of view of land-in-general, economic rent is a surplus, not a cost. The rent of a particular piece of ground depends, first upon the general level of rents, and secondly upon its differential advantages of fertility, situation, climate and so on, over other pieces of land; from the point of view of a single branch of agriculture, all or part of this rent is a cost, not a surplus.

**Demand.**

Given the supply of any factor of production, its earnings depend upon the demand for its services. Here the governing influence is productivity—the more productive a factor is, the greater will be the demand for it, and the higher it will be paid. Productivity might seem to imply the creation of something of real social value, but as we have seen, the term is used in economics in a much narrower sense to refer to the creation of utility. Society might be little the worse if expert manicurists and animal dentists were exterminated, and a great deal the worse if farmers and bricklayers disappeared. But manicurists may nevertheless be paid more than farmers because they are more "productive"—more productive in the sense that a greater value is set on their services by those who have plenty of money with which to back their valuations. If there were as many manicurists as farmers, however, it would be the farmers who were more productive. We would have more manicurists than we could possibly use, and they would require to offer their services for next to nothing. Just as the utility of a commodity declines the more we have of it, so the productivity of any factor of production declines the more of it we are already employing. There is a law of diminishing productivity parallel to the law of diminishing utility.

**Marginal Productivity.**

It is, in fact, not simply the productivity of a factor that governs the price which we will offer for its services (our "demand price"). It is the productivity of the factor at the margin, or its marginal productivity. The earnings of any factor of production tend to be equal to the value of the marginal product of the factor. This may be defined as the value of the contribution to output made by that unit of the factor which is engaged in the least productive task.

Suppose, for example, that a farmer is hiring workers all of whom are of equal efficiency and all of whom, if competition is perfect, will have to be paid the same wage. It may be worth £5 a week to the farmer to have the services of one ploughman, £3 a week to have the services of a second, and £2 a week to have the services of a third. The second ploughman will not be so indispensable as the first: if, for example, the farmer has no difficulty in laying down part of his land to grass but wishes to keep some minimum area under the plough; or if there are enough odd jobs about the farm to keep one ploughman busy in slack times but hardly enough for two. Similarly, the work done by the third ploughman will be less vital to the farm as an enterprise than the work of the second ploughman. There may also be other workers on the farm—cattlemen, shepherds, etc.—who are entrusted with tasks of varying productivity but who are all on the same level of efficiency. If the wage at which agricultural workers can be hired is fixed, the farmer will add to the number of men whom he is employing until the addition to the produce of the farm made by the last or marginal man just balances his wage. To employ more men would mean a needless sacrifice of profit; to employ fewer would be to miss an opportunity of a further small profit. Thus the marginal man will earn a wage equal to the value of the tasks which he performs, and other workers, being of the same efficiency as the marginal worker, will be paid the same wage whatever the value of the tasks on which they are engaged. If the workers are all interchangeable so that the farmer is quite indifferent which of them he employs, he will have no reason to pay a high wage to one man and a lower wage to another. He will pay all alike the value of the marginal product of agricultural labour.

What is true of one farm will be true of all. Each farmer will take on more men up to the point at which the marginal product of labour is equal to the wage that has to be paid. If, when all farmers are trying to do this, there is a shortage of labour, then farmers on whose land the marginal product of labour is above its wage will require to bid away workers from farms on which these workers are not being employed to the best advantage. This will force up wages; a variety of tasks which were previously undertaken by labour will cease to be worth while; the demand for labour will be reduced; and equilibrium between demand and supply will be restored at a higher level of wages and marginal productivity. On the other hand, if wages are so high that farmers do not find it profitable to employ all the workers who are looking for jobs in agriculture, it will be necessary for wages to fall, so extending the field of employment to include tasks which were previously submarginal, before the unemployed workers can find jobs.[1]

[1] This argument must not be taken to imply that if there is heavy unemployment the proper remedy is to cut wages. Heavy unemployment usually occurs during a trade depression and can best be cured by measures calculated to end the depression. It is doubtful whether wage-cuts can be included amongst such measures. See pages 412-3.

The marginal product of labour may be greater on one farm than on another. While competition for workers is making itself felt, we may find a prosperous and expanding farm on which the marginal product of labour is above the level of wages, and a depressed and contracting farm on which the marginal product of labour is below the level of wages. The first farm will be trying to increase its staff, the second to diminish its staff. But if competition is effective we will not find the first farm paying high wages and the second low wages. The state of profits on any single farm will not, so long as labour can move freely from one farm to another, affect the terms on which that particular farm can hire its labour. Nor will wages be high or low on any single farm because the marginal product of labour *on that farm* is high or low. We must measure the marginal product of labour not on a farm which is more prosperous or more depressed than the general run of farms, but on a representative farm which will reflect the trend of farming. Such a farm will be in equilibrium when farming is in equilibrium, and will make excessive or inadequate profits according as the industry is progressing or depressed. The rate at which a representative farm hires its workers will be the rate which other farms also will require to pay. Agricultural wages are governed, on the side of demand, by the marginal product of labour on a representative farm. This conclusion applies not only to agriculture, but to all industries. For "farm" we can just as well read "firm." In each industry wages tend to equality with the marginal product of labour in a representative firm.

The line of reasoning which we have so far followed can readily be extended to cover the whole of industry. If there is free movement of labour between industries—and this is far less likely than free movement between firms—then the wages paid to labour of a given grade of skill and efficiency will be the same in each industry. The state of prosperity of an industry will not affect the terms on which it hires its workers. Those terms will be governed by the marginal productivity of labour in a representative industry which is sensitive to the dominant trends in employment, expanding or contracting as the general level of employment expands or contracts.

**Marginal Productivity and Substitution.**

When we studied the demand for a commodity we found the idea of the margin bound up with that of substitution. In studying the demand for a factor of production we find this association repeated. The rate at which the marginal productivity of a factor diminishes depends upon what we might call its "substitutability"—that is, upon the ease with which additional units of the factor can be substituted for other factors. If two factors can be readily substituted for one another, the marginal productivity of either of them will diminish slowly, for there will be a large number of tasks performed by one factor which could be easily undertaken by the other if more of it were to become available. Now our ability to substitute one factor for another is limited, as we have seen, by the law of

diminishing returns. This law is little more than a statement that the substitution of one factor for another becomes progressively more difficult as the ratio of combination alters. It is a law, so to speak, of diminishing "substitutability." In the law of diminishing returns, therefore, we have one reason—a technical reason—why marginal productivity declines. Additional supplies of any factor can be absorbed into employment with increasing difficulty, because, in the extra uses to which the factor is put, it is a less and less effective substitute for the other (constant) factors. Its value in these extra uses—the value of its marginal product—diminishes, and the price which it can command falls correspondingly.

Diminishing marginal productivity can also be traced to the law of diminishing utility. If any factor is less scarce, its products will also be less scarce and, therefore, of less utility. They will fetch a lower price, and the factor which produces them will be forced to content itself with lower earnings. Here we have a second reason—a psychological one—why marginal productivity declines.

If the supply of any factor of production increases, therefore, there will be a double substitution—by producers, in conformity with the law of diminishing returns, and by consumers, in conformity with the law of diminishing utility. Suppose, for example, that more people want to find jobs in agriculture. Then increased competition will force down agricultural wages until farmers have an incentive to take on more men. This incentive will be twofold.

First, since labour is cheaper relatively to capital and land, it will pay, in the growing of any particular crop, to use more labour and less capital and land. Farmers who were forced by the high cost of labour to instal milking machines, and to use tractors for ploughing, will now be able to go back to hand-milking and horse-ploughing. Where cultivation was extensive, so as to economise labour, it will become more intensive as the use of labour becomes less costly. Thus a given quantity of product will be raised by a larger number of men, working either with less capital or on a smaller area of land than before.

Secondly, the fall in agricultural wages will not affect all branches of agriculture equally. Wages are a much larger item in the cost of wheat than in the cost of wool, and they are a still larger item in the cost of potatoes and strawberries. A fall in wages, therefore, will reduce the price of wheat relatively to the price of wool, and will reduce the price of potatoes and strawberries even more. The greater the reduction in price, the greater will be the increase in demand; consumers will buy more of the relatively cheaper and fewer of the relatively dearer goods. There will be a great expansion in strawberry growing and those branches of agriculture which make most use of labour, and a comparatively slight expansion in sheep-farming and those branches of agriculture in which labour is least used. Consumers will substitute goods using much, for goods using little labour and the total demand for labour will, therefore, increase, The more readily consumers can adjust their budget so as to transfer demand from one group of goods to the other, when wage-costs are

rising or falling, the more sensitive will be the demand for labour to a change in wages; the more elastic, in other words, will be the demand for labour.

**Elasticity.**

The rate at which the marginal productivity of a factor diminishes as the supply of the factor increases is the same thing, if the factor always earns the value of its marginal product, as the elasticity of demand for the factor. Thus our conclusions as to diminishing marginal productivity can be easily re-stated in terms of elasticity—the elasticity of demand for any factor will be governed by the technical obstacles to substituting that factor for other factors of production, and by the psychological obstacles to substituting its products for their products. It will depend also upon the elasticity of supply of appropriate substitutes, both to producers and to consumers.

Although each factor of production can be substituted for the other factors at the margin, substitution cannot be carried so far that we can do without one factor altogether. It is difficult to imagine a type of production into which land, labour and capital do not all enter. Thus the factors are simultaneously in competition and in co-operation with one another; they are at once rival and complementary. This two-sided relationship, which arises because the factors are simultaneously in joint demand and in composite supply, raised problems similar to those touched on in Chapter 17. How, for example, will a change in the supply of one factor react on the earnings of the others?

Suppose, first of all, that there are only two factors of production—labour and property. Then an increase in the supply of one will necessarily raise the earnings of the other. If more workers have to be employed, the marginal productivity of labour will fall, and each worker, not just each *additional* worker, will earn a lower wage. Since the additional men employed will be contributing to the value of output at least as much as their wages, the residue of product after payment of wages to all workers, will be greater. This residue will be equal to the earnings of property. The return on property will necessarily increase. But the *share* of property in the total product may not increase. If the demand for labour is elastic, employers will spend more on labour as it becomes cheaper, so that a greater aggregate wage-bill will be distributed amongst the larger number of workers. The aggregate earnings, both of labour and of property, will increase; and if the demand for labour is sufficiently elastic—in other words, if labour is easily substituted for property—it will be the earnings of labour that increase most. The wage-bill will form a higher proportion of the National Income, although wages per man will, of course, be lower.

There is reason to believe that labour and property are not *in fact* easily substituted for one another. If we think of property as made up of machinery, this conclusion will seem very puzzling; machinery

is constantly being substituted for labour, and if labour became cheaper instead of dearer the substitution might be reversed. But if we think of property as made up of dwelling-houses, power stations, railways, agricultural land, and so on, our difficulty will be to see how any kind of substitution of labour is possible; substitution has to be more roundabout—as, for example, scarcity of domestic servants induces us to move from an old house that needs a great many servants to run it, to an expensively fitted flat that needs very little domestic help. Since property of the second type predominates over property of the first type, labour and property are not very good substitutes for one another. They are more complementary than rival, and an increase in one, therefore, tends to diminish its share in the National Income, and to increase the share of the other.[1]

When there are more factors than two, an increase in the supply of one factor may be of advantage to some and of disadvantage to others. If, for example, America were to admit a large number of Chinese immigrants, the wages of domestic servants, dock labourers, laundry workers, etc., would probably fall, while the middle classes would find themselves better off because they would be able to buy the services of Chinese workers more cheaply without suffering any loss of income either from property or from work. The immigrants would be almost entirely rival to the first group of workers, and almost entirely complementary to the second. There might also be an intermediate group of workers, such as farmers or semi-skilled factory workers who, on balance, would neither gain nor lose appreciably.

The conclusions reached above are very far-reaching in their implications. For example, it will be in the interest of any group of people that workers in other trades should be as numerous and hard-working and efficient as possible, while other workers in their own trade, or in close competition with it, should be as few, as lazy and as inefficient as possible. The first condition will make for cheapness in what they buy and the second for dearness in what they sell. Again, it will be in the interest of the working-class that capitalists should accumulate property as rapidly as possible rather than dissipate their income in private extravagance, since as the stock of property increases the return on it will fall, the earnings of labour will rise, and the share of labour in the National Income will very probably rise also. On the other hand, it will be in the interest of the propertied classes if the population increases rapidly, if hours of work remain long, and if wage earners work hard and save little. It is clear that the analysis of this chapter provides a clue to many serious conflicts of interest between classes and countries.

[1] Over the last hundred years, property has been increasing more rapidly than labour and yet the share of each in the National Income has been remarkably stable. Presumably invention has been maintaining the marginal productivity of property just sufficiently to offset the influence of its growing abundance.

**Marginal Product and Marginal Net Product.**

The marginal productivity theory, as stated so far, is based upon an assumption which we have not yet investigated. That assumption is that it is possible to *measure* the productivity of each factor of production. But if labour is always used in conjunction with capital and land, how are we to disentangle the peculiar contribution of labour from the joint product of all three? How much of the timber cut by a lumberman is to be "attributed" to his labour, how much to the capital sunk in his axe, and how much to the natural fertility of the soil? According to the marginal productivity theory, this question can be answered only by going to the margin and measuring the value of the timber cut by an additional lumberman. But if he has no axe he can cut no timber, and his productivity will be nil. If he *has* an axe, we have still to decide what *its* productivity is. We seem to be no further on. The theory of marginal productivity ceases to be applicable if an increased quantity of one factor can be used only in conjunction with a correspondingly increased quantity of the other factors. In other words, the theory presupposes that it is technically possible to vary the proportions in which the factors of production are combined.[1]

This is not a very extravagant assumption. If methods of production were dictated everywhere by technique, the task of the employer would be an easy one. No search after improved methods, no substitution of one factor for another in a ceaseless effort to keep down costs! Instead, the automatic adoption of routine technique. unchanged from boom to slump and back again! In fact, however, employers are not so fortunate. Technique is plastic and responsive to changes in the relative scarcity and price of the factors of production. Variations in "technical coefficients"—the ratios in which the factors of production are combined—can be made in three ways.

First, employers may use extra units of one factor with a constant quantity of the others by changing the form in which the other factors are provided. They may, for example, employ more lumbermen and cheaper axes so that everyone has an axe but not such a good one as before. In the average industrial concern, even a change of this kind would not be necessary; it would be sufficient to institute a night-shift with the existing plant.

Second, there are generally wide variations in technical methods between the firms in an industry, some firms relying on a lavish use of machinery and plant, and others on a lavish use of labour. These variations are possible because of differences in the scale of operations and in management. But if the cost of either labour or capital increases, some firms will suffer a disproportionate reduction in profits because of their dependence on the dearer factor; and these firms will tend to be squeezed out of production, while the firm which employ a larger proportion of the other factor will continue.

[1] For the similar problem of joint cost, see page 204.

Third, the prosperity of different industries equally with the prosperity of competing firms, depends upon the comparative cost of the preponderant factors of production. If the course of wages is upwards, the coal industry, which employs a very high proportion of labour to capital, will be more embarrassed than the chemical industries in which capital preponderates. It is possible to vary the range of uses for labour while the supply of other factors remains constant by expanding one set of industries and contracting another.

Since these variations in technical coefficients do take place it is unnecessary to assume that employers can or should attempt to measure the marginal productivity of each factor. It will be sufficient if they measure something rather different—something which, in equilibrium, will be equal to marginal productivity. This alternative concept is marginal *net* productivity.

The marginal net product of any factor is the net value of the contribution to output made by a marginal unit of the factor when combined with the appropriate quantity of other factors; a "marginal unit" being *any* unit of the factor when engaged in the least productive use to which it is put; and "net value" being understood to mean the total value less the cost of extra units of the other factors. Thus the marginal net product of lumbermen will be the value of the timber cut by the marginal lumberman, less the cost of his axe and the royalty on the timber. *Provided* these last items can be taken for granted, it will be enough to think in terms of marginal *net* product. The marginal net product of lumbermen will govern the demand for their services and will tend—given perfect competition—to be equal to their wage.

**Marginal Productivity and Surpluses.**

It might appear that if one factor of production earned no more than its value in a marginal use, other factors of production would necessarily reap a surplus income; or that, if all the factors of production earned the value of their marginal product and no more, their joint income would fall short of the value of their total product. We have already seen, however, that under perfect competition price is equal to average cost, including normal profit; there is therefore no surplus income accruing as profit to the entrepreneur. Price is also equal to marginal cost, and the value of a marginal addition to output is equal therefore to the joint earnings of marginal units of each of the factors, combined in the optimum ratio to one another. This is just another way of saying that each factor is paid the value of its marginal net product. But it implies the absence of any surplus. For since each unit of output has the same value and requires just as much effort to produce, and since units of the factors of production are paid alike, what is true at the margin is true of total output and total earnings. The social product is exhausted when each factor earns the value of its marginal net product, and in conditions of perfect competition this will be equal to the value of its marginal product.

**Long-run Influences on Earnings.**

In order to give concreteness to the general theory outlined above, we may try to list the various influences which, in the long run, will make for an increase in the earnings of one of the factors of production. For this purpose we may concentrate on labour. In what circumstances should we anticipate a general rise in wage-earnings?

(i) Labour may become scarcer. This will restrict the range of tasks on which labour can be employed, eliminate tasks of inferior productivity, and so raise the productivity of labour at the margin. Or at least this is what would happen given constant returns. But if it requires a large population to exploit advances in technique—if, for example, a railway system, hydro-electric power stations, and so on would not come into existence without it—then the answer is not quite so simple.

(ii) Land and capital may become less scarce or it may become possible to make more intensive use of the existing supply (e.g., by growing two crops where one was grown before or by running two shifts instead of one) This will reduce the marginal productivity of land and capital and induce employers to compete more eagerly for the limited supply of labour. There will be a larger *number* of tasks to be performed in conjunction with the additional supply of other factors, and it will be necessary, therefore, to reserve labour for the *kind of task* in which other factors cannot readily be substituted for it. The use of labour will be limited to tasks of higher productivity than before and the marginal productivity of labour will, therefore, increase. For example, if farmers suddenly found themselves with larger areas of land to cultivate, they would probably have to take men from work on the farm which just repaid itself and set them to plough the extra land. The work of ploughing being of greater productivity than the work from which the men were taken, farmers would be able to pay higher wages, and if they wished to retain the services of their workers in face of competition from other farmers, would be *forced* to pay these higher wages.

(iii) Workers may work harder, or faster, or for longer hours than before The supply of *work*, although not of *labour*, will increase, the range of tasks performed by labour will be extended, and the marginal productivity of work will be reduced. Workers will be paid less per unit of effort; but, if the demand for work is elastic, the total reward for effort (the wage-*bill*) will increase, and since this total has to be divided amongst a constant number of workers, wages also will increase. A greater disposition to effort will be to the advantage of the other factors of production in the same way as an increase in the number of workers. From the point of view of the other factors it makes little difference whether labour is more abundant or more hard-working; either change increases the supply of work. It is only from the point of view of labour that there is a difference. If there are more workers, each will be paid less; if everyone works harder, and the demand for work is elastic, each worker will be

paid more. Conversely, if the demand for work is inelastic, workers will be able to earn higher wages by showing a greater disinclination to effort. The other factors of production will stand to lose from such a restriction of the supply of work just as they stand to lose by a reduction in the number of workers.

(iv) Workers may become more efficient, or better trained, or more skilful. Since these changes increase the productivity of labour, each of them will, in general, lead to a rise in wages. An increase in training or in skill will *always* tend to raise wages, but an increase in efficiency may conceivably operate to reduce wages. Faster work, as we have seen, will reduce wages when the demand for effort is inelastic; and an increased speed of working is just as likely to result from greater efficiency as from greater application. Some *kinds* of increase in efficiency, however, are quite certain to raise wages. Greater efficiency in the use of property, for example, will reduce the number of breakages, fires, etc., and allow the supply of property to increase more rapidly. Greater efficiency in the use of loan-capital, by directing it into industries where it is less likely to be wasted, will have the same effect. Unless the property which is "salvaged" in this way is mainly of the type which is in close competition with labour (e.g., machinery), labour will necessarily gain from the greater abundance of property.

(v) The productivity of factors other than labour may increase; for example, new inventions may be made. How do inventions affect wages? Do they make for a rising standard of living for everyone, or do they lead to unemployment and low wages?

**Invention and Wages.**

Inventions are generally "labour-saving"—that is, they generally economise labour more than they economise capital or land. The greater the economy of labour, the stronger will be the tendency to unemployment and lower wages. But an invention may be labour-saving and yet increase both employment and wages. The railway engine, for example, has probably done more to save human effort than any other single invention. Yet nothing did more to increase the volume of employment in the nineteenth century than railway-building, and nothing contributed more to the raising of the standard of life than cheap railway transport. Similarly, although labour-saving machinery generally creates unemployment and low wages in the industry into which it is introduced, workers in other industries may derive a more than counterbalancing advantage from the fall in costs and prices to which the invention leads. So far as can be judged from the history of labour-saving inventions, they have generally done much more to raise than to lower the general level of wages, both by improving the worker's standard of living and by making labour less irksome. But if they have been of advantage to labour, they have been of even greater advantage to capital, since they have increased the productivity of capital more than the productivity of labour, and so have increased the *share* of the National Income accruing to

capital.[1] It should be observed, too, that labour-saving inventions are not of equal advantage or disadvantage to all classes of workers. An invention may, for example, be skill-saving or toil-saving. If the invention is skill-saving, and strikes at an established handicraft such as handloom weaving, the earnings of skilled workers as a body will tend to suffer. If, on the other hand, the invention is toil-saving, and like the mechanical navvy economises muscular effort, then it will be the earnings of unskilled workers that tend to be reduced. In either event, since greater mechanisation is involved, machine-minders are likely to benefit.

Some inventions are not labour-saving, but capital-saving or land-saving. Communication by wireless, for example, requires less capital than communication by telephone or by deep sea cable. In urban transport, the motor omnibus is less capitalistic than the Tube. Similarly a new fertiliser which increases the fertility of the soil economises land, while urban sites are economised by improvements in the construction of skyscrapers. These inventions, while they may increase the total earnings of capital and land, just as labour-saving inventions may increase real wages, will tend to reduce the *share* of capital and land in the National Income. Possibly because it has been labour which, of all the factors of production, has risen most in price over the past two hundred years, most inventions have been labour-saving, while capital- and land-saving inventions have been comparatively rare. On the whole, therefore, invention has been of greater service to property than to labour, and has done more to raise the return on property than to raise wages.[2]

**Elements of Monopoly.**

We can now relax the drastic assumption of perfect competition which has been made throughout this chapter. Employers do not bid for more labour, land and capital without regard to the effect of their bids on the price at which they hire the factors of production. They do not exclude, in deciding what to pay and whom to employ, everything except their immediate financial interest. And they are rarely able to strike bargains with individual wage-earners as to the rate of wages which they will pay. In most industries wages are fixed by a process of collective bargaining between employers' associations and trade unions; sometimes by direct negotiations, sometimes by arbitration after argument of the case by both sides before an independent chairman; sometimes by an intermediate

[1] If, for example, the National Income is divided in the ratio of 55 : 45 between labour and capital, and labour-saving inventions increase the National Income by 20 per cent., the enlarged National Income may be divided between labour and capital in the ratio of 60 : 60.

[2] It is perhaps unnecessary to point out that just as some workers gain while others lose by an invention, so some property-owners will gain while others lose. There is an obvious loss to those whose property becomes obsolete, and an equally obvious gain to those who own the invention or who use its products.

procedure of conciliation and discussion between the parties to the dispute with an added representation of independent or semi-independent opinion.

Collective bargaining, whatever the form which it assumes, is bargaining in which monopoly power is exerted by both sides; competition continues to influence the outcome, but the full force of competition is not exerted. The maximum rate which employers will pay is governed by the range of alternatives open to them; the minimum rate which the factors will accept is governed by the range of alternatives open to them. But the alternatives vary with the degree of competition. Under perfect competition no employer would offer a wage higher than the value of the marginal product of labour since, if wages exceeded marginal productivity, the dismissal of workers would yield an increased profit; and no worker would accept less than the value of his marginal product since if this were refused him by one employer he could find work at a higher rate of wages elsewhere. Immediately collective bargaining takes the place of private bargaining between each worker and each employer, the alternatives are transformed. There is still a maximum which employers will offer—the loss which they would suffer through the withdrawal of labour. There is also a minimum which workers will accept—what they could earn in another occupation less the cost of changing their occupation and the money-measure of their preference for their existing job over other work. But between the demand-price of employers and the supply-price of the workers there may be a wide range of indeterminacy, within which bargaining power has free play.

On the tactics of the trials of strength that occur in wage disputes something will be said in the next chapter. For the present, only the broad strategy is in question.

**Demand Price in Conditions of Monopoly.**

The rate of wages which employers are willing to offer falls short of the value of the marginal product of labour for a variety of reasons.

In the first place, employers are not disposing of their output in conditions of perfect competition. They tend therefore to restrict their output, and simultaneously their demand for labour, so as to maintain equality between marginal revenue and marginal cost. If they lose or dismiss a worker, they sacrifice, not the full value of labour's marginal product, in the sense of physical product multiplied by price, but only the marginal physical product multiplied by marginal revenue. The output of an additional worker may be worth £5 a week on the assumption that additional sales are effected at unchanged prices; but if prices have to be lowered, the net return to the employer will be less than £5. The more monopolistic the market for a given product, therefore, the lower will be the demand price for labour in relation to its marginal productivity, and the more restricted will be the employment of labour.

In the second place, even if competition in goods were perfect, employers have an incentive to limit their demand price for labour whenever they feel that increased employment will involve them in paying higher wages. If the value of the marginal product of labour is £5 employers may continue to pay £4 and abstain from taking on more men because of the higher rate of wages that they would have to pay, not only to the extra men but also to those already in their employment. Under collective bargaining, this source of difference between marginal productivity and demand price tends to disappear, since each employer is free to vary his labour force at the standard rate of wages; in technical language, the supply of labour to any one employer has become perfectly elastic where before it was slightly inelastic. But a large employer may still feel that if he were to bid for all the labour that it would be worth his while to employ at the current rate of wages, he would be in danger of precipitating new claims for wage advances. The State might, for these reasons, pursue a more modest building programme than it would otherwise contemplate. In general, however, it is only in the absence of collective bargaining that employers will limit their demand for labour for fear of raising wages.

In this they will be fortified by custom and regard for other employers. Farmers, for example, will generally limit their offers of wages to some customary rate, partly from self-interest but partly so as to stand well with their neighbours. In domestic service the same is often true; some people will go without domestic help rather than pay much more than the customary rate.

**Supply-Price in Conditions of Monopoly.**

When wage rates are dependent upon the outcome of collective bargaining, there is less room for personal effort to secure better wages by moving from one employer to another. Each employer has to pay the same standard rate of wages, and if this falls below what the worker is prepared to accept he must either remain unemployed or move to some other industry. The supply price of labour to any one employer is no longer the rate paid by his competitors, but depends primarily on the rates paid by other industries in the district or (in some circumstances) on the rate of benefit paid by the Unemployment Assistance Board.

**Monopolistic Influences on the Earnings of Capital.**

It is no more true of capital than of labour that its earnings are fixed by perfect competition. One business has easy access to capital because of the personal contacts of its directors, while another, almost equally credit-worthy, is forced to expand out of profits. Interlocking directorates, the ramifications of holding companies, customary association, and all the links, tangible and intangible, that bind one business to another serve to break up the capital market into semi-monopolistic groups In countries with slender resources of capital a few financial institutions may be almost

the sole channels through which domestic savings (and still more, borrowings from abroad) can flow into investment. These institutions will obviously be immensely powerful, and the terms upon which they lend are not likely to be those upon which the same supply of capital would be offered by a new banking house seeking only to cover its costs. In Great Britain the general abundance of capital limits the power of the banks and finance houses, but in backward countries the danger of monopolistic exploitation is a very real one.

**Unemployment.**

The fact of unemployment takes greatly from the value of the marginal productivity theory. For the theory assumes throughout that all of the factors of production will be absorbed into continuous full employment, and that rates of earnings must be such as to bring this about. If, in fact, earnings remain steady and it is employment that fluctuates, then the theory of marginal productivity misconceives the problem and we must try another line of approach.

Over very long periods there can be no doubt that rates of earnings serve to equilibrate the supply of, and demand for, the factors of production; and the assumption of no unemployment does not destroy the usefulness of the marginal productivity theory as a guide to the response of earnings to changes either in supply or in demand. But, in the short or fairly short period, fluctuations in employment, and the persistence of involuntary unemployment, must be accounted for, and their influence on the distribution of income incorporated in the general theory.

No attempt will be made at this stage to expand the theory of distribution to cover fluctuations in employment. But an illustration may be given of the way in which a too hasty application of the marginal productivity theory leads to a quite false deduction. It would seem to be an obvious corollary of the theory that a reduction in wage rates will produce an increase in employment. But this is true only of a reduction in *real* wage rates. If money wages are reduced, as we shall see,[1] the cost of living will probably fall simultaneously, leaving real wages as they were. That is, it is a great deal more difficult than one might suppose to make real wage rates fall, and the apparently obvious cure for unemployment of cutting money wages may not achieve the intended result.

**Summary.**

The argument of this chapter has been conducted at a high level of abstraction from real life. But the problem of distribution is so complex that it must be approached in stages, and the first stage is to understand the tendency of competition. Setting aside the influence of monopoly and unemployment, we have seen that the distribution of the National Income between the factors of production is governed by the relative scarcity of the factors and by their marginal productivity.

[1] See pages 412-3.

The supply of labour and capital depends upon our marginal preference for work and thrift, while the supply of land is fixed by nature. The ratios in which the factors are combined are determined by employers, who give effect to the twin principles of substitution and diminishing marginal productivity, varying their demand for any one factor until its marginal productivity is equal to its price or earnings. The removal of the assumption of perfect competition introduces an element of indeterminacy. The earnings of a factor will lie between the lower limit set by the supply price of the factor and the upper limit set by its loss value to employers.

## CHAPTER 20

## WAGES

### § 1

THE theory of wages deals mainly with two questions—what determines the general level of wages and the share of labour in the national product? and why do wages differ in different places and occupations? An answer to the first question has already been given in the previous chapter. There it was assumed that all wage-earners were alike in efficiency, training and skill so that no employer would offer higher wages to one worker than to another. It was assumed also that all wage-earners were alike by temperament and taste, and that they worked under similar conditions in occupations and places that were equally agreeable or disagreeable, so that no worker would demand higher wages in one job than in another. On these assumptions we found that the tendency of competition was to make the general level of wages equal to the value of the marginal product of labour.

The second question arises as soon as we drop the assumptions. Wage-earners differ in efficiency, training and skill, and these differences are reflected in differences in wage rates. But even amongst workers of the same general calibre there are fairly large differences in wages. Why is this? Why is it that workers who are free to choose their occupation and place of work continue to accept low wages when they might move to better-paid jobs? Why do different industries and districts pay different rates of wages for the same grade of work?

**Nominal and Real Wages.**

First of all, differences in wage rates are generally expressed in terms of the full-time weekly rate of wages—sometimes called "nominal wages." But the nominal wages payable in any occupation may be greater or less than the average income which workers in that occupation can normally expect to earn. The cash wage may be supplemented by an allowance of free coal, or potatoes, or house-room,

and these perquisites—which are of particular importance in the earnings of agricultural workers—obviously form part of the " real wages " payable to the worker. On the other hand, a deduction should be made from nominal wages in some occupations to cover any expenses on uniforms, overalls, travelling expenses, social display, etc., associated with the worker's duties. The bank clerk who is dragooned into the respectability of bowler hat, white collar and tailored suit may properly deduct part of the cost of this uniform in arriving at his real wage. Nominal wages will also differ from real wages if overtime, short time or unemployment are normal features of an occupation. A bricklayer whose work is frequently interrupted by bad weather will not be so well off as a postman whose nominal wages are the same and who has the advantage of steady employment. On the other hand, if there is a great demand for houses and a shortage of bricklayers, it will be the bricklayer, not the postman, who earns most. As for unemployment, an occupation like shipbuilding which suffers from periodic slumps will require to offer fairly high nominal wages before the real wages obtainable in it over a period of years reach a moderate level. A final point of distinction is that whereas nominal wages are expressed in terms of money, real wages consist of the goods and services which this money will buy. Thus if we find a 'bus driver in London earning £5 a week and a 'bus driver in Glasgow earning only £3 a week we must not jump to the conclusion that wages are higher in London than in Glasgow, for it may cost £5 to buy in London what can be had in Glasgow for £3. Differences in nominal wage between occupations and places may be quite consistent, therefore, with equality of real wages.

**Net Advantages.**

Secondly, workers may accept comparatively low wages in one place or occupation because of some compensating advantage which they enjoy over places and occupations where wages are higher. There may be opportunities, for example, of alternative or supplementary employment either for the worker himself or for his family. If an engineer loses his job, or is dissatisfied with it, he has a better chance of finding work with another firm if he lives in an industrial district like Glasgow instead of in a Highland village. He will be inclined, therefore, to demand a higher rate of pay from an electric supply company in Fort William than from a similar company in Glasgow. A textile worker may prefer to accept low wages in Lancashire, where all his family can find employment, rather than move to a well-paid job in Lincoln or Shrewsbury, where his family might earn little or nothing. Similarly, a married man with a young family may hesitate to give up agricultural work because of the ease with which he can find work for his children on the land. If he happens to have a great deal of spare time every winter, and can use this spare time for toy-making, his income from this supplementary work will be an added inducement to remain in agriculture.

Opportunities of rapid promotion and prospects of spectacular earnings will also compensate for comparatively low initial earnings, and this compensation will be greater than, on a strictly rational view, it should be. Since most of us have a good conceit of ourselves, and exaggerate our own chances of success, we allow ourselves to be unduly influenced by the *maximum* earnings of an occupation, and are not always deterred by low *average* earnings. The knowledge that the most eminent lawyers earn more than the most eminent doctors will have more weight with an ambitious young man than the knowledge that the average doctor earns more than the average lawyer.

Compensation for a comparatively low wage may also be found in various amenities which attach to the job and enhance its attractiveness —what might be called the "kudos other than cash"—which the worker enjoys. A man may prefer to seek employment in a poorly-paid job because the sacrifice in earnings is more than balanced by these non-pecuniary amenities. He may find the work more agreeable because it is carried on in the country, or in a well-ventilated factory, or with less danger to health. He may be attracted by the regularity and security of the job, or by the opportunities of leisure or absenteeism which it affords, or by the power and prestige attaching to it, or by the scope which it offers for the exercise of some special bent. On the other hand, he will avoid work which is exhausting (e.g., mining), or causes great discomfort (e.g., road-breaking), or endangers life (e.g., the work of the steeplejack, the sailor, etc.), or which is carried on in an unhealthy atmosphere (e.g., the roasting of copper ore, stoking, the manufacture of chemicals, etc.), or which is monotonous or excessively speeded up.

What determines the attractiveness of a trade, in other words, is not just the wage which it pays but its Net Advantages. Even when workers are alike in efficiency, and can move freely from one job to another, competition will not drive wages to the same level in all places and occupations. It will be the Net Advantages of each place and occupation that tend to equality. There are, however, a number of obstacles to the free movement of labour and hence to the equalisation of Net Advantages. These obstacles, and the resulting immobility of labour, provide a third reason for difference in wages between places and occupations.

**Mobility and Immobility.**

When an industry becomes comparatively overcrowded, as coal-mining did after 1921, it is not possible for workers to transfer themselves immediately to other industries. Movement would mean the sacrifice of an acquired technique, and a long and irksome period of re-training. The older the worker, the greater the difficulty. The older men are less adaptable, and have more at stake in the revival of the industry. In it their industrial experience has taken shape, and in it they have become accustomed to ways of life which they are reluctant to abandon. They are the more willing to wait in the

hope of re-employment because their chances of employment in other industries are small. Employers will hesitate to start them on comparatively unfamiliar work so long as there is an over-supply of younger workers in the same district. The burden of change, therefore, falls on these younger workers. The same is true of geographical movement. If there is a depression in one part of the country, it will be particularly difficult for the older workers to move in search of employment. They are tied by family responsibilities and by old associations. They may own their house and be forced to sell it at a bargain price. On the other hand, they may have great difficulty in renting or buying a house in districts where there is no lack of work. The mere cost of transferring a family from one district to another may be enough to deter a married man from moving to a better-paid job. The young worker, with no family and no property, will move much more readily.

Thus changes in the distribution of labour between places and occupations come about more through the action of young workers in moving where prospects are good than through transfers on the part of older workers out of contracting industries and localities. Young workers are not yet in a groove. They have fewer commitments and more enterprise, and do not need to overcome the inertia of an earlier choice. The higher the proportion of workers in the younger age-groups, therefore, the greater is likely to be the mobility of labour. Since the proportion has in fact been declining for over forty years, and will almost certainly continue to decline for some time, there are clearly strong forces at work making for a less mobile and less adaptable population.

Even the mobility of young workers, however, tends to be rather low. The range of occupations open to them, particularly in districts like South Wales, and towns like Bradford, is generally narrow, and the real range of choice is often even more restricted by family circumstances. The need of working-class families for immediate income may force children into blind alley occupations. Or, by the force of example of parents and relatives, children may be induced to enter the industry with which their family is most familiar. This may, and does, happen even when the industry is in process of contraction and the older members of the family cannot find work in it. The tendency for boys to follow their fathers' footsteps, even when the footsteps go deeper and deeper in the mud of unemployment, is not due solely to bias towards the parental occupation, but quite as much to limited knowledge and limited contacts. Neither the boy nor his parents are likely to be so knowledgeable about alternative occupations, and the father may lack both the means and the influence to give the boy a start in some other industry. Juvenile labour, therefore, does not always flow to the points of scarcity, but tends to remain within grooves cut by family circumstances.

Immobility may result from the worker's disinclination to move—a disinclination originating in ignorance, uncertainty, poverty, family ties, or sheer inertia. But there are other, external, reasons for

immobility; there may be impediments to the free movement of labour. A man cannot leave one country and enter another whenever he hears that wages are higher abroad;[2] there are laws restricting immigration.[3] Even within a country his movements may be restricted —an agricultural worker may require permission to take a job in industry, a teacher born in one county may be debarred from employment in other counties, or there may be discrimination by race, language or religion, limiting certain occupations to members of a single group. When such restrictions have not the force of law, prejudice, bias and custom may give rise to a boycott which is no less effective in limiting the field of employment. There are many occupations, for example, from which women have never been excluded, by law, but in which, until recently, they rarely if ever found employment. This has partly been due to the pressure of trade unions and the rules of professional associations like the General Medical Council. These bodies, in an effort to increase their powers of monopoly, have not unnaturally sought to limit their numbers by erecting barriers against entry into their occupation or profession. Women—amongst others—have suffered since they have been forced to squeeze into occupations where entry is freer and which, as a result, are comparatively overcrowded. One reason, therefore, why women's earnings are lower than men's is just that they are not allowed to participate in the monopoly gains of "protected" industries, but have to crowd into the limited range of occupations which men have thought fit to surrender to them.

**"Unfair" Wages.**

Immobility of labour, whether due to disinclination to move or to artificial impediments to movement, creates disparity between the Net Advantages of different places and occupations. Labour of the same calibre will be paid higher wages or enjoy better conditions of work in one place or occupation than in another. A comparative surplus of labour in one trade will reduce its marginal productivity there, while a comparative shortage in another will raise the marginal productivity of labour there. Wages in the two trades will diverge and the divergence will continue until there is a movement of labour out of the first trade and into the second. If, for example, the demand for coal falls, then coal-miners will probably be paid lower wages and will obtain less regular employment than before. Coal-mining will be comparatively unattractive, but miners will have to put up with the change in conditions because of their inability to find work in other industries.

The net advantages of coal-mining will be less than the net advantages of other occupations employing workers of equal efficiency. This is sometimes expressed by saying that the wages of miners will be "unfair," the term "unfair" being used to mean "less than can be earned with equal effort by similar workers in other industries." In this sense, the wages of agricultural workers

are almost invariably "unfair." As the world becomes richer a smaller proportion of its income goes on necessaries and a larger proportion on luxuries. The higher the standard of living, therefore, the smaller the proportion of our productive resources that is required for the production of foodstuffs, most of which can be classified as necessaries. We need fewer farmers and more doctors, engineers and so on. But it is exceedingly difficult to induce workers on the land to transfer their efforts to some other line of production. Since their children are numerous and are generously biased towards agricultural work, the tendency is for a larger, not a smaller, proportion of the labour supply to seek employment in agriculture. This attachment to a constantly declining industry makes it necessary to squeeze out farm workers by paying them "unfair" wages. So long as agriculture is overcrowded, foodstuffs have to be sold at low prices, and the incomes both of farmers and of agricultural labourers are depressed below the level which equal skill and effort would procure in other industries. These low incomes, moreover, are earned by all farm workers, not just by those who will ultimately be forced to give up farming. Thus the relative over-supply of farming labour exposes it to what is, in effect, exploitation by workers in other trades who get their food at abnormally low prices. The immobility of farm workers keeps down their wages to an "unfair" level.

In the world of the future a high degree of labour mobility will be more than ever essential, partly to secure speed of adjustment to violent changes in technique and in the direction of trade, and partly because there will be in the labour market a higher proportion of older and less mobile workers. In the last twenty years innovations in methods of manufacture have been at a prodigious pace, which shows no sign of slackening; great industries have toppled headlong and younger industries have thrust their way upwards; and at the same time the age-distribution of the working population has been less and less favourable to the rapid displacement of old methods and industries by new. In the next twenty years these things will undoubtedly continue. The more we can increase mobility, therefore, the better.

The methods by which this can be done are well known—the provision of fuller and more accurate information to workers in search of alternative employment (e.g., through the Labour Exchanges); grants towards the cost of movement (e.g., through schemes of assisted transference); the opening of Training Centres for unemployed workers or workers in overcrowded industries. But the familiar methods must be used imaginatively and on a sufficient scale. It was a mistake, in the years before the war, to restrict the Government Training Centres to unemployed men; they should be open to everyone, just as a secondary school education should be open to everyone. It was also a mistake to hope for a high degree of mobility so long as there were few jobs to move to; there was little incentive to attend a Training Centre so long as those who did remained out of work like the rest

of the unemployed. If we want high mobility we must get rid of unemployment.

So far, we have taken a kind of horizontal cross-section of labour and confined the discussion to workers of equal efficiency, training and skill. But when we think of differences in earnings we generally have in mind inequalities whose origin lies in " vertical " rather than in " horizontal " differences between workers. We think of lawyers and film-stars and company directors and compare their earnings with those of clerks and charwomen and navvies. We reflect that, far from these differences in earnings being compensated by differences in amenities, it seems to be the best-paid jobs that require least effort and afford most leisure. We may even ask ourselves why, if the jobs at the top are so attractive, so many people remain at the bottom. Three reasons suggest themselves—lack of ability, lack of training and lack of opportunity.

Differences between workers in natural ability are exactly like differences between pieces of land in natural fertility; just as the more fertile pieces of land command a premium over the less fertile, so the more able workers can obtain higher wages than the average. Indeed, economists often speak of a " rent of ability " when comparing the earnings of superior talent with the earnings of mediocrity, so extending to labour the conception of rent as a differential surplus. This surplus, like land rent, is determined solely by demand. It is not enough to have rare gifts, superior to the gifts of others, in order to earn an exceptionally large income. It is necessary also to have those particular gifts which are most in demand. Few artists have ever made a fortune, or even a decent livelihood; just as few landlords have ever made money by improving the landscape. The ability of a statesman, a social reformer, or a philosopher is far more poorly rewarded—except in prestige—than the ability to cheat and dissimulate and take advantage of the ignorance of others. The most " able " workers—like the most " fertile " pieces of land—are those who can provide the services for which people pay, not those who, on a moral or æsthetic judgment, are most deserving of reward.

Differences in training also produce differences in earnings. If a worker has to sink capital in a long and expensive education (e.g., at a university) or in acquiring skill (e.g., in an apprenticeship), he will expect to be recompensed by appointment to a post which is either more agreeable or better paid than the posts open to him without training. A barrister, for example, has to spend many years, during which he earns little or nothing, preparing for examinations which he may never pass, and waiting for briefs which he may never be given. He has to pay stiff fees to examining bodies and to professional associations, and may also have a premium to pay on apprenticeship. Only the prospect of a high income—in cash or in amenities—will induce him to make the necessary sacrifices; he will look to be compensated, in the course of his career, for all his outlay, and for the long period of waiting and uncertainty—to say nothing of the cost of midnight oil.

So far as the cost of training is part of the supply price of labour—that is, so far as higher earnings are a necessary inducement to training—the extra earnings of trained workers will contain no element of rent corresponding to the rent of superior ability. But if at any time there is a shortage of trained workers, and apprentices cannot be trained quickly to meet the demand, the limited supply of trained workers will be able to obtain abnormally high earnings. Their earnings, in the short period during which no more workers can be trained, will be governed solely by demand and will be swollen by an element of *quasi-rent*. Quasi-rent resembles rent in that it is a surplus over cost or supply price, but differs from rent in that it disappears in the long run when supply is given time to increase.[1] Compare, for example, the earnings of a professor of mathematics and those of a 'bus-driver. The professor can go on earning a high salary because few people are capable of expounding Einstein; but the 'bus-driver can rarely earn much more than an unskilled worker, since there are large numbers of unskilled workers who could fairly easily learn 'bus-driving. The high salary of the professor contains an element of rent—rent of ability. The 'bus-driver's wage, on the other hand, will contain an element of rent only if, during a temporary shortage of drivers, their wages have been increased. This element of rent will be transitory—a *quasi*-rent—since in course of time more trained 'bus-drivers will become available.

Inequality of earnings may also be due to inequality of opportunity. Workers are not equally lucky in meeting with opportunities of earning higher wages. In particular, they are not equally lucky in their choice of parents. Some people are given a better start in life than others because of the superiority of their parents in character, or in education, or in wealth. The children of rich parents, for example, enjoy a more expensive education and a wider cultural background, have more freedom in their choice of occupation and better guidance in choosing it, and have more opportunity of making valuable social contacts. Thus although they may ultimately be outstripped by the abler members of a lower social class—a class with more limited opportunities—they start with a fairly large advantage in income-earning power, and are generally able, given moderate ability, to keep ahead in the high-income group. Social mobility, that is, is generally low. A worker tends to stay in the class into which he was born. He is unlikely to slip downwards because of the advantages of education and training which he enjoys, and he requires more than average ability and determination to fight his way upwards under the handicap of limited opportunities.

Just as there are often artificial impediments to movement between places and occupations, so there are often artificial impediments to movement between social classes. Under the caste system, for

[1] This use of the term "quasi-rent" is to be carefully distinguished from its use in the following chapter to mean "the earnings of an instrument of production." On this second definition, the rent of a house is a quasi-rent.

example, social mobility is nil. In countries inhabited by different races the dominant race often monopolises positions of responsibility so that members of the subject-race cannot climb to a higher status. Even in Britain it is not unknown for a man to be excluded from an important post simply on the grounds that he is the son of a manual labourer; or for applications for highly-paid jobs to be restricted to public schoolboys. The effect of these restrictions is to shelter a privileged class of persons from competition and to swell their incomes with a surplus or rent, akin to the surplus profits of a monopolistic producer. A somewhat similar surplus emerges when social mobility is limited by inequality of opportunity rather than of privilege. The differential advantages of birth in a rich family are as substantial, if not as secure, as the differential advantages of birth in a superior caste. Advantages of either kind are the product of institutions devised by man—the institution of inheritance or of the caste system—and the surplus income or rent corresponding to these advantages is called, therefore, "institutional rent." Since human institutions can be modified by appropriate legislation (e.g., by laws abolishing caste or limiting bequest) this source of inequality can be reduced at will (i.e., at the will of those who command political power). But it can never be entirely eliminated. It is impossible to contrive that everyone shall have exactly equal opportunities. We cannot, for example, equalise the ability of parents to bring up children, although we can see to it that the children are given more or less equal opportunities of education in schools and universities. We can increase social mobility, but we cannot make it perfect.

**Summary.** On a horizontal plane competition makes for equality of net advantages between places and occupations, while immobility —either geographical or occupational—limits the effectiveness of competition and preserves inequalities. Immobility may be reinforced by restrictions on movement which buttress the monopolistic advantages of particular groups of workers. On a vertical plane, competition tends to equate earnings from work with innate differences in efficiency (as assessed by employers or by consumers) or with differences in efficiency which are the product of education and training. Competition, however, takes place within an institutional framework which limits social mobility by creating inequalities of opportunity. These inequalities tend to restrict access to highly-paid occupations, which are recruited mainly from the well-to-do, and the consequent failure to make full use of the supply of innate ability in other classes of the community makes for excessive differences in earnings. Inequality of earnings from work, therefore, is partly due to immobility and partly to differences in efficiency. Hence if we wish to reduce inequality, we must either increase mobility or abandon efficiency as a criterion of reward.

§ 2

The rate of wages in any industry, then, depends upon the competitive pull which other industries are able to exert on its labour, upon the

cost of moving to the most attractive alternative industry, and upon the strength of the workers' disinclination to move—that is, it depends upon the supply price of labour to the industry. It goes without saying that the rate of wages depends also upon the demand price of employers. For workers of given productivity this will be governed by the price (or, under imperfect competition, by the marginal revenue) of the product of the industry. In the short run, fluctuations in wage-rates are closely linked with fluctuations in prices; the most powerful force bringing about an increase or decrease in the rate of wages paid to any group of workers is an increase or decrease in the price of the goods which they make. In the long run, when there is time for a change in the relative attractiveness of an industry to cause workers to move from it or to it, it is the supply price of labour that is the dominant factor; much as, in the long run, it is to cost rather than marginal utility that prices respond.

The rate of wages in an industry depends also upon a third factor—the strength of the organisations of workers and employers. In the trials of strength to determine where, within the limits of indeterminacy, wage rates are to be fixed, it is open to employers' associations to threaten to lock out their workers and to trade unions to retaliate by the threat of a strike.

The position of an Employers' Association is generally a strong one. Where there are few masters and many men it is easier for the masters to hold together than for the men; it is easier for them to frame a common policy, and they can negotiate with more assurance and authority. They have the advantages of education and influence; they can put their case more skilfully before the public—or even before the men; and they can, if they choose, bring pressure to obtain publicity of the right sort. Above all, they have greater reserves. Especially in a depression, the prospect of closing down for a time while surplus stocks are disposed of, and the organisation and plant are overhauled, is not altogether unattractive. But for the worker his wage-earnings are normally his sole source of income, and strike pay out of the accumulated funds of his union is poor compensation. The longer he is idle, the greater his poverty and the less his ability to sustain himself in a fit condition for advantageous employment. Thus while employers may declare a lockout merely to improve their prospects of future profit, it generally requires an angry resentment or an obstinate devotion to some cherished view of right and wrong to provoke the mass of wage-earners to a lengthy strike.[1]

[1] Not that most strikes are about wages—far from it. Strikes because of working conditions, dismissals, allegations of victimisation, or to secure recognition for a Union, are much more frequent although generally less prolonged. There are also plenty of strikes for no reason at all except a deep feeling of frustration: violent outbursts of insubordination and disaffection.

There is, in short, a fundamental disparity between the bargaining power of master and man. The onus is on the worker to sell his labour to procure his family's daily bread, but there is no comparable obligation on the employer to provide work. The employer can nearly always be sure of replacing a worker, but the worker cannot be sure of finding another employer quickly.

This disparity is softened by non-economic influences. The employer (particularly in an established firm) has a reputation to preserve and a pride in his business. It would be a travesty of the facts to suggest that in the middle of a wage dispute employers pack their bags and go off for a round of golf by the seaside while their workers brood over an empty larder and their common wrongs. Employers may be under less compulsion to re-start, but are generally no less eager. Moreover, although it ought to be easy for them to make common cause, they have a way of falling out with one another and playing blackleg. Their calling puts a premium on masterfulness and private judgment and releases them from much of the discipline of joint effort; there are often some wayward spirits, therefore, who refusé to abide by their association's "official" plan of campaign.

**The Economy of High Wages.**

There are also economic factors disposing employers to pay higher rates of wages than they need. The economic interest of the employer is not in low wages but in low *wage-costs*. He pays a certain rate of money-wages and receives in exchange the services of his work-people. The more favourable the terms upon which he can hire these services, the greater will be his profits. This does not mean, however, that it will pay him to offer low rates of wages; low-paid labour is often the dearest labour.

First of all, low wages mean a low standard of living, and this in turn means a low standard of efficiency. Workers who are not properly fed and housed, or who are left with little leisure, are unlikely to reach the outputs of more highly-paid labour. They cannot stand up to the strain of modern industry so readily, or show the same initiative, or undertake as responsible work. When we find that workers in India and China are paid much less than British workers, therefore, we must not jump to the conclusion that wage-costs are correspondingly low, and that competition with such countries is impossible. The fact that India is still one of Britain's biggest markets suggests that low wages have their disadvantages. On the other hand, we must not jump to the topsy-turvy conclusion that India would compete more successfully if employers there were to grant immediate increases in wages. Workers do not begin to work harder whenever their wages go up. Their first reaction is often to relax and either work *less* hard, so as to keep their earnings at much the same level as before, or spend less carefully, using the rise in their pay haphazardly and without advantage to their efficiency. It is some time before higher wages seep through to raise the level of efficiency, and since

employers can rarely afford to take the long view—they may lose their workers at any time to their competitors—they are unlikely to pay high wages with an eye solely to their workers' standard of living. Particularly if this standard is already high, any gain in efficiency will appear small and speculative while the cost of a rise in wages will be large and certain.

There are, however, two further motives which are likely to be stronger in their appeal to the self-interest of the employer. The first is that by paying high wages it may be possible to "cream" the labour market—that is, to attract workers of more than average efficiency. The second is that high wages may purchase the goodwill of a firm's employees, and create an atmosphere in which they are not deterred by suspicion and discontent from working wholeheartedly on behalf of the firm. Both of these motives, it will be observed, may induce a firm to keep wages high in relation to its immediate competitors in the district from which its labour is drawn; neither motive is sufficient to ensure that wages will be high *throughout* an industry or district.

**Trade Unions.**

The self-interest of employers is hardly the best guarantee of high wages. In order to secure improvements in their standard of living and conditions of employment, workers generally prefer to rely on their own organisations. The chief aim of these is to press for higher wages, shorter hours, and better factory amenities. Sometimes trade unions have more militant objects, such as the overthrow of the capitalist system by direct action. But in this country the trade unions have about as much revolutionary ardour as the Post Office, and like the Post Office are too busy collecting funds for provident benefits (e.g., sickness and unemployment), to contemplate a *coup d'état*.

Apart from any specific functions, one of the prime objects of trade unions has always been to give the worker a sense of status—the feeling that he is one of a body of workers organised to protect their rights and interests, not alone and defenceless against the whims of employers. Again, trade unions, by employing skilled negotiators, can put the workers' case with more finesse, confidence and knowledge and can take his part in the hundred and one day-to-day grievances on things not covered by any formal contract. This is not without its advantages to the employer also, for it creates a regular channel through which grumbles can reach him, and allows him to enter into agreements which are binding on all his workers. The danger of irresponsible strikes is, therefore, reduced although by no means removed.

From the point of view of the community, too, trade unions serve a useful purpose in preventing employers from resorting in bad times to the attractive device of cutting wages as a means of cutting costs; and in bringing pressure on employers to examine carefully fresh methods of economising labour. But the pressure of trade unions does not always make for increased efficiency. In the first place,

when unions are organised on a craft (or occupational) rather than on an industrial basis, they are apt to have narrow demarcation rules and insist that given jobs should be done by members of their craft even when they involve little skill and could be done with much more convenience by other people. The cumulative effect of these rules in holding up urgent work and in perpetuating obsolete industrial practices is to cause a great deal of waste of time and labour contrary to the social interest, although very much to the monopoly advantage of the union. Rules limiting the number of learners and apprentices to be employed with each journeyman, or the number of women to each man, may have a similar effect. These rules are often no more than a safeguard against attempts by employers to get work done cheaply by learners who are not given an adequate training, and for whom no provision is made by the employer at the end of their time. But the tendency of such rules is plainly monopolistic and they are generally pushed beyond the point at which they are in the social interest. Finally, if trade unions stick out for high wages over the whole field of industry, they may restrict the volume of employment exactly as a monopolist restricts his output of goods, and may achieve a high standard of living for those who continue in employment, quite as much at the expense of those who fail to get employment as at the expense of employers as a class.

**Limits to the Bargaining Power of Trade Unions.**

Fear of creating general unemployment will not deter a particular trade union from pressing for higher wages. Trade unions in expanding industries will not regard it as part of their duty to keep wages low so as to ensure the rapid absorption of unemployed labour from less fortunate industries. Even the fear of unemployment for their own members is less compelling when rates of unemployment benefit begin to be comparable with wage-earnings.

The limits to the power of trade unions to raise wages in a single industry are threefold :—

**(i) Elasticity of Substitution.**

The more readily employers can economise labour by using alternative methods of production, the more effective will be their resistance to claims for wage advances. The simplest expedient is to instal more machinery ; employers may re-vamp the layout of their plant so as to replace old-fashioned machines by more modern labour-saving types, or they may be induced, even in up-to-date plants, to turn over to more highly mechanised processes which were just not worth while when labour was cheaper. Either way, capital will be substituted for labour by a change in the proportion in which the two are used. The rate at which substitution proceeds—the rate at which the proportion of the two factors changes as their relative price changes—is called the elasticity of substitution ; and the elasticity of substitution is the first damper on wages.

Substitution is not confined to the replacement of labour by capital; replacement of labour by labour, directly or indirectly, is quite as effective. If employers can find "blacklegs," or import labour from other areas, or get women or unskilled workers to do what was previously done by skilled men, they need trouble themselves no further about substitution. Alternatively, they may replace workers in one occupation by workers in another. Rather than pay higher wages to ploughmen, for example, farmers will tend to substitute cattlemen and shepherds by putting more land under stock and less under crops. In the same way, if stonemasons extort a thumping increase in wages, and bricklayers and carpenters are paid at the same rates as before, builders will tend to put up more houses of brick or wood and fewer houses of stone.

**(ii) Elasticity of Supply of Alternative Factors.**

How far substitution can be pushed depends not only upon the ease with which technical changes can be made, but also upon the ease with which additional supplies of the substitute can be procured. In a boom or in war-time, machine tool manufacturers may be booked up for months ahead and practically all workers absorbed into employment. These are obviously favourable times for seeking wage advances. On the other hand, substitutes are in highly elastic supply during a depression and the bargaining power of trade unions is then at its lowest.

**(iii) Elasticity of Demand.**

Consumers as well as producers can indulge in substitution, and the measure of their ability to find substitutes is the elasticity of demand for the product of the industry. Where demand is elastic, employers have difficulty in passing on an increase in wage-costs to consumers and they tend, therefore, to offer more resistance to demands for higher wages. A case in point is provided by the export industries which supply a fiercely competitive market and in which, between the two wars, even the strongest unions were hard put to it to maintain wages while the trend of wages in other industries was upwards.

These limitations of the bargaining power of trade unions are, in effect, limitations on the demand price for labour and are tantamount, therefore, to a threat of unemployment if wages are pushed too high. Apart from this ultimate sanction, there is generally a good deal of "play" in wage-rates depending on the current fortunes of the industry. If, for example, profits are high, and employers enjoy monopolistic or semi-monopolistic advantages, labour may be able to obtain a share in the swag and secure rates of pay which are well above the general run of wages in other industries. The employers may even go out of their way to make concessions to their workers in order to disarm criticism of their monopoly. A different kind of blackmail may be practised by a small but truculent union in an occupation which forms an indispensable part of large enterprises

(e.g., the printing of newspapers). Rather than have the whole of the work held up from time to time by demands for higher wages from a relatively insignificant number of workers, employers may seek to forestall these demands and treat the extra outlay in wages as a cost normal in their line of business.

**What do Wage Disputes really settle ?**

If the workers in an industry secure an increase in their wages and do exactly the same work as before, who is it that surrenders the income that the workers gain ? That someone must be worse off is plain ; the National Income is no greater, and one section of workers is receiving more. Who, then, is " squeezed " ?

It is common to assume that the extra wages come out of profits and that it is the employers, therefore, who are the losers. But if employers can pass on the extra burden of cost in the form of higher prices, they can maintain their profits and it is then the consumer, not the employer, who is " squeezed." Now " the consumer," when we are dealing with any large industry, includes other groups of workers as well as those who spend out of profits or rent. The gains of one section of workers, therefore, are likely to be in part at least at the expense of other workers. If wages rise all round, the real income of labour may be no greater than before, for the cost of living may have risen *pari passu*.

This is not simply an academic contingency but what commonly happens. Twice in war-time we have seen violent increases in wages and prices, and on both occasions the two have kept closely in step with one another. It is true that this was at least as much because wages followed prices upwards as because prices responded to the upward trend in wages ; but the fact is striking none the less. In the famous Blum experiment in France in 1936 it was only too obvious that the general rise in wages was the controlling factor in the corresponding rise in prices. The experience of the United States under the National Recovery Administration in 1933 points to the same conclusion.

The fact is that wage disputes have comparatively little effect on the distribution of the National Income between capital and labour ; they are much more likely to settle the distribution of the wage-bill between one section of labour and another, or between one industry and another. In the last analysis an increase in wages in one industry represents a levy on other industries and on the general consumer. If they cannot stint their requirements of the goods made by the industry, consumers will have to meet the whole of the increase in wages in the form of higher prices. But if demand is elastic, some part of the burden will recoil on employers in reduced profits and on the workers in reduced employment.

Broadly speaking, therefore, wage disputes are disputes over *relative* rates of wages. They do not fix the return on capital. But if not, what does ? To this question we must now turn.

# CHAPTER 21

# INTEREST AND PROFIT

## 1. INTEREST

INTEREST is the price paid for the hire of loan-capital; more briefly, it is the price of a loan. This price is usually expressed as an annual rate, calculated on the principal of the loan. If, for example, I borrow £100 for one year on impeccable security, I am likely to be asked to pay back £105 at the end of the year—that is, interest will be charged at the rate of 5 per cent. per annum. But why should I be asked to pay interest? What has my creditor done to earn his £5 in interest?

**Why Interest is Paid:—**

**(1) The Supply of Capital.**—The immediate reason why interest is paid is that loan-capital is scarce. The amount of money which people are willing to lend every year falls far short of the amount which would be borrowed if no interest were charged. So it is necessary to ration out the limited supply by putting a price on loans—that is, by charging interest on them. Only those borrowers who are willing to pay the current rate of interest will be able to obtain loans; those who cannot pay the price will be forced to go unsatisfied.

But why is loan-capital scarce? Why don't people lend more? There are three possible reasons:—

(*a*) Unless we are bankers, we cannot lend without going to the trouble of saving, and there is a limit to our willingness and ability to save. We can, however, lend to any single borrower without saving more; all we need do is to ask for repayment from some other borrower and lend the sum repaid.

(*b*) We cannot lend without denying ourselves the use of our own savings. If we save more, therefore, we do not necessarily lend more; we may, instead, apply our savings to increase our stock of goods. For example, we may buy machinery for use in our business; or make a speculative purchase of raw materials to guard against, or profit from, a rise in their price; or store foodstuffs in fear of war; or buy a house, or a motor-car, or some similar durable commodity. In other words, we may choose to lend to ourselves rather than to other people, and hold our savings in the form of *goods* rather than in the form of debt.

(*c*) We cannot lend without parting with our money for a period of time and leaving ourselves in a comparatively illiquid position during that time. Money is more convenient for many purposes than even the best of I.O.U.s, and because of this greater convenience we are never willing to lend out all our money. In other words, we may prefer to keep part of our savings in the form of *money* rather than in the form of debt.

Of these three reasons we shall, for the present, disregard the last, and treat the second as a variant of the first. We shall assume, that is, that when people save more they will also lend more. On this assumption, the only possible reason for the scarcity of loan-capital is the scarcity of savings. Upon what, then, do savings depend ?

**(i) The Rate of Interest.**—It might be supposed that the supply of savings, like the supply of oranges, depended mainly on their price, i.e., on the rate of interest. It is generally agreed, however, that although the payment of interest does offer an inducement to thrift, this inducement is subordinate to others. A low price would make us give up orange-growing, but it might require a *negative* rate of interest to make us give up saving. We don't grow oranges for lack of anything better to do with our time ; our *sole* motive is likely to be the profit which we hope to earn. But we may save money because we can find nothing better to do with our income ; millionaires, for example, will go on saving automatically whether the rate of interest is high or low. If the price of oranges rises from 1d. to 1½d. we can be quite certain that more oranges will be put on the market. But if the rate of interest rises from 3 per cent. to 4 per cent. we cannot be certain that savings will increase. A man who wishes to save just enough to yield him £300 a year in interest when he retires will save *less* at the higher rate ; at 4 per cent. he will require to accumulate only £7,500, whereas at 3 per cent. he will need £10,000. Similarly, life insurance companies may find that their premium income (which is a substantial part of the total savings of the community) is reduced rather than increased by a rise in interest rates. It is doubtful, however, whether many people have such rigid ideas about the future income at which they are aiming, and it is likely that, on balance, a rise in interest rates will increase savings. This conclusion is reinforced by a further consideration. A rise in interest rates redistributes the income of the community in favour of the rentier class. Now rentiers, being rich, save a higher proportion of their income than the rest of the community. The changed distribution of income, therefore, will tend to increase total savings. Here there is no question, however, of a greater *incentive* to thrift. If more is saved it is because capital happens to be concentrated in the ownership of a small group of rich people instead of being distributed equally or monopolised by the State.

At very low rates of interest there would probably be a heavy falling-off in savings, while at very high rates savings might be greatly stimulated. A rate of 2 per cent., for example, would tempt many people to consume their capital by purchasing an annuity instead of keeping their capital intact and trying to live on the interest. On the other hand, rates of 8 or even 10 per cent.—such as governments might be driven to offer in war-time—would put such an enormous bounty on saving that many people would make sacrifices in their habitual standards of living in order to add to their future income.

Within the range of interest rates to which we are accustomed, however—from about 3 to about 5 per cent.—it is doubtful whether saving shows much elasticity.

The nominal rate of interest is not an accurate measure of the *effective* inducement to thrift. First we must deduct from the nominal rate of interest any taxes payable by the lender. If there is an income tax of 5s. in the £1, I must pay this 5s. whether I spend the £1 or save it. But if I save it, I shall later be asked to pay income tax on the interest. When the nominal rate of interest is 4 per cent., therefore, the real rate to the saver is only 3 per cent. His choice lies between spending £100 this year and having £104 less £1 in tax next year. On each £100 saved he gains only £3. Secondly, we must make provision for changes in the value of money. If I expect prices to rise by 5 per cent. over the year, I shall make a bad bargain by lending money at 4 per cent., for £104 in a year's time will buy less than £100 does now. The *real* rate of interest will be *minus* 1 per cent. On the other hand, if I expect prices to fall by 5 per cent., and my expectations are well founded, the real rate of interest will be 4 plus 5, or 9 per cent. A given rate of interest, therefore, offers a greater incentive to thrift in times of falling than in times of rising prices. If we wished the real rate of interest to remain constant, we should require to lower the nominal rate when prices were falling and raise it again when prices were rising. Strangely enough, this is exactly how the rate of interest does behave, although the changes in the rate of interest are generally small in comparison with the changes in the value of money. The reason for these sympathetic movements in interest rates and prices is not, as one might suppose, that more is saved when prices are falling, nor even that there is an increased pressure to lend, but rather that people are less willing to borrow at rates which are *nominally* low but *really* (in relation to price-trends) abnormally high.

**(ii) Social Institutions.**—Savings depend upon the encouragement and reassurance offered to the saver by existing social arrangements and practices. If there are numerous outlets for savings, if lenders are brought readily into contact with borrowers, and if there is widespread approval of thrift as a civic virtue, the disposition to save will be immensely stronger than in communities where capital is little in demand, the capital market is inefficiently organised, and the general temper of society is unfavourable to thrift. Thrift is promoted also by a high code of business morality which gives savers confidence in the good faith of their debtors and permits loans to be made without undue fear of default; and by a tranquil and orderly social system which protects life and property and allows savers to reap the fruits of their thrift. Habits of saving, therefore, vary from one society to another. They were stronger under Queen Victoria than under Queen Elizabeth, and a great deal stronger than under Queen Boadicea. They are stronger in a stable capitalist society than in communities where the danger of repudiation or confiscation

is serious, or where to possess capital is to court persecution or murder.

**(iii) Income.**—Every increase in our income adds to our ability to save. Of course, we are unlikely to save the whole of the increase, for when we become richer we generally scale up our standard of living. But neither are we likely to spend the whole of the increase. Some part will be added to our savings—especially at first, while our customary standard of living is still unadjusted to the change in our income. Experience suggests that this part is generally fairly large, and that we tend to save not only a larger sum of money but also a larger *proportion* of our income. Experience suggests also that if, instead of comparing the same person at two levels of income, we compare rich people and poor people, rich people as a class save a larger proportion of their income than poor people as a class, and that the richer the class, the higher is the proportion of income saved. It follows that a rich community saves more than a poor community, even when the poor community is exceedingly thrifty. The richer the community, the greater is the danger that saving will be carried to excess.

**(iv) Wealth.**—Our ability to save depends also on our wealth. The more wealth we have, the greater is the income which it yields and the more readily, therefore, can we add to our accumulated wealth by further saving. Wealth grows like a snowball ; as has often been remarked, it is easier to make the second million than the first hundred. A poor man is likely to have less capital at eighty than he had at sixty ; but a rich man, if he gives nothing away, will probably have twice as much.[1] If wealth is inherited, the cumulative advantages of ownership can be transmitted and perpetuated. Inequalities originating in differences in ability or in luck crystallise into class differences. Saving tends to become the prerogative of the wealthy, and the wealthy become entrenched as a class behind the barrier of inherited property.

An increase in our wealth enables us to save more only if there is an accompanying increase in our income. An increase in wealth unaccompanied by any increase in income will dispose us to save *less* of our income, since the more wealth we have the weaker is the motive to further accumulation. If our holdings of Stock Exchange securities rise in price, for example, we are likely to have less hesitation in spending out of our current income ; while, if we feel ourselves impoverished by a fall in security values, we may spend our income less freely. In the United States consumer-spending is remarkably sensitive to any major fluctuation on Wall Street.

**(v) Thrift.**—In a given order of society, at a given level of income, how much we save depends on the strength of the motives disposing

[1] *cf.* J. C. Wedgwood ; "Economics of Inheritance" (Pelican edition), page 191.

us to thrift. These motives rarely induce us to make provision for savings by setting aside a definite sum; unless, of course, we are under contract to save—for example, by making repayment on a mortgage to a building society, or by paying an annual premium on our insurance policy. These payments apart, our expenditure is governed largely by the desire to maintain some customary standard of living. If our income is more than sufficient to maintain this standard, we save the surplus; if it is less than sufficient, we borrow the deficit (e.g., by running up debts to tradesmen or an overdraft at the bank) or live on our capital. In the short run, therefore savings are a residue after meeting customary expenses. In the long run, however, people have standards of thrift as well as of expenditure. If they find that they are saving consistently more than they wish, they will raise their standard of living. On the other hand, if they find that they are running into debt, or are saving less than they think desirable, they will try to economise and reduce their expenditure. In the long run, savings cease to be residual and are fixed more and more deliberately.

Our standards of thrift vary with our preference for future as compared with present goods. If the rate of interest is 5 per cent., we can choose between £100 now and £105 next year. Those of us who are thrifty will have a comparatively strong preference for £105 next year. We may be uncertain of the future and anxious to provide for emergencies; or hopeful of achieving a greater measure of independence and a higher standard of living; or eager to have more capital at our disposal to expand our business and back speculative projects; or full of ambition to build up a fortune and acquire social status; or we may have more foresight or be more alive than others to future needs and future pleasures; or we may want to accumulate money for its own sake out of sheer miserliness. Other people, less concerned for the future, or attaching less importance to the possession of wealth, will discount future goods more heavily and be more attracted by £100 now. Everyone, whatever his motives to thrift, will modify his preference for future goods the larger the proportion of his income which he saves until at the margin he is indifferent whether his last pound is spent or left to accumulate interest—that is, until his marginal preference for future goods is measured by the rate of interest.

Much of the savings of wage-earners and salaried workers are designed to provide for future needs or to supplement future income. Such savings have been described by the late Lord Stamp as "long-distance spending." A man may save in order to buy a motor-car or to pay for a holiday. He may build up a reserve to support him in sickness or old age, or to meet the cost of educating his children. This kind of saving, in spite of the increasing provision for social insurance which is made by modern governments, appears to be greater than ever. It would seem that social insurance, by underpinning the standard of living, has made thrift seem less futile and its fruits more secure and attractive.

**Corporate Saving.**

Nearly half the savings of a modern industrial community are made by companies, not by individuals. Out of their total profits, companies distribute part in dividends and "plough back" the remainder into the business. This "ploughing back" of profits springs from much the same motives as induce individuals to save for speculative projects which they have in view. The chief motives are enterprise and prudence. A business which is making large profits will generally be tempted to expand further, and it will be much less troublesome to finance expansion out of profits than to make a new issue on the Stock Exchange. Since directors are frequently given credit for rising dividends which are the product of reinvested profits rather than of increased efficiency, the allocation made to reserve may be more generous than enterprise alone would justify. A second slice of profits may be saved in order to provide the business with liquid reserves against such contingencies as increased competition, strikes, industrial depression, and so on, or in order to equalise dividend payments over a period of years. There may be some savings also which are included in the allowance made for depreciation of plant, if this allowance is inflated from motives of prudence; for the sinking fund out of which the plant will ultimately be renewed will then grow more rapidly than the plant wears out, leaving a residue of savings when renewal does take place.

**Government Saving.**

An increasing amount of saving is being undertaken by local authorities and by the central government. While the bulk of the capital outlay of governments is financed out of loans, provision is generally made for extinguishing the loans out of a sinking fund, and these governmental sinking funds now form a large part of our total savings. Much of the ordinary expenditure of government bodies, moreover, is designed to improve our stock of human or of material capital. If part of the tax revenue of the country goes to meet the cost of education, or of health services, or even of unemployment insurance and relief, then although no physical assets come into existence, workers benefit exactly as they would by drawing on their savings to improve their education, or to maintain their health when ill or their efficiency when out of work. If grants are made for the construction of roads, hospitals and public buildings, these grants may also be reckoned in the total savings of the country

**The Justification of Interest.**

We have seen that, for a variety of reasons, savings are scarce and command a price which we call the rate of interest. But does this *justify* the payment of interest? Can we, for example, regard interest as compensation for the sacrifice involved in saving as we can regard wages as compensation for the irksomeness of work?

Like the earnings of any other factor of production, interest is both a price and a source of income. Regarded as a price, interest

performs a useful social function by rationing out a limited supply of savings between competing borrowers. Even in a socialist community, this is a function which interest might continue to perform, for it would still be necessary to determine the priority of schemes calling for the use of capital and to eliminate those schemes which showed an insufficient return. If the return could be measured in money terms, the charging of a rate of interest would automatically secure this result, since it would give priority to schemes offering prospects of a higher return, and weed out schemes of lesser productivity. The rate of interest would be fixed just low enough to attract claimants for the total supply of savings. Regarded as a source of income, interest is less easy to justify. It is paid to some people rather than to others, because some people own capital while others do not. If we wish to justify interest *payments*, therefore, we must first justify these differences in ownership. Secondly, we must show that without the offer of interest, capital would cease to be accumulated or cease to be lent. The first point raises issues too wide for adequate treatment here. Differences in ownership are an integral part of the capitalist system. Until we have considered the merits of capitalism, therefore, it is difficult to decide on the merits of one of its chief by-products—inequality. For we may judge that inequality is not too high a price to pay for the advantages listed by apologists for capitalism; the premium which it puts on efficiency; the scope which it offers to men of initiative to back their judgment and foresight against the opinion of their competitors; the incentives which it offers to the accumulation of capital in a world which looks on capital accumulation as one of the surest tests of progress. We may believe—as many economists of the last century did believe—that nothing could so effectively harness men's energies for the satisfaction of their mutual wants as the prospect of unequal gains and the free enjoyment of property rights. Even the most convinced defender of capitalism, however, might be willing to grant that where inequality is the fruit, not of effort, but of inheritance, fraud or luck, it is unjustifiable and ought to be diminished by such expedients (e.g., Estate Duties) as leave the fundamental institutions of capitalism undisturbed.

The second point raises quite different issues. We know that it is only at the margin that the rate of interest has a decisive influence on thrift, and that the vast bulk of savings would continue to be supplied at lower rates. Interest, that is, generally contains an element of rent—a saver's surplus in excess of the sacrifice to which the saver is put. We know also that the stock of *existing* capital puts its owners to no sacrifice at all. It is already there, just as land is already there. True, capitalists may *consume* part of their savings (or the savings of past generations) at low rates of interest, so that interest payments may be a bribe to dissuade them from this alternative. Except at very low rates, however, capital consumption would probably be negligible. Long before the sources of thrift dried up, or the stock of capital was drawn on, a more immediate danger would make itself felt—the disinclination of capitalists to *lend*.

**Time-Preference and Liquidity-Preference.**

This brings us to the second reason why loan-capital is scarce. We have to overcome not only our reluctance to save, but also our reluctance to lend. Just as we may prefer present to future goods, and must be offered interest in order to overcome our "time-preference," so we may prefer money now to the promise of money at some future date, and require the offer of interest before we can overcome our "liquidity-preference." The phrase "liquidity-preference" sums up the various motives which prompt people to hold money rather than lend it at interest. The more we prefer money to I.O.U.s of any kind—i.e., the greater our liquidity-preference—the more difficult it will be to induce us to lend, and the higher, therefore, will be the rate of interest.

The conception of liquidity-preference is one which cannot easily be understood without some explanation of what money is and what it exists to do. This explanation we must postpone. We can, however, go part of the way towards an understanding of liquidity-preference by asking why this phrase is used to describe our demand for money. What is this mysterious thing "liquidity" and what connection has it with money? By "liquidity" is meant "power to convert into other commodities or into something which is generally acceptable in final settlement of a debt." Now money is of all things the most readily exchangeable into other commodities and the most generally acceptable in payment of a debt. Nothing is easy to get rid of in all circumstances. Nothing, that is, is more liquid. Money, therefore, sets a standard of liquidity against which we can measure the liquidity of everything else. So the term "liquidity" comes to mean "power to convert into money." A preference for liquidity is a preference for money.

Consider, for example, the liquidity of bank advances. A banker who makes a loan to a merchant for the purpose of buying stocks of goods can rely on rapid and more or less automatic repayment, even in a crisis, because stocks are comparatively liquid; they can be readily converted into money so long as consumers go on buying, and will be converted automatically as the merchant disposes of them in the ordinary course of trade. Loans of this kind are liquid, therefore, because the stocks which provide security for the loan are liquid. On the other hand, an advance to a householder to enable him to buy his house will not be particularly liquid. The householder will not be put in possession of funds with which to repay the loan by merely living in the house. In a crisis he can pay off the loan in cash only out of his savings (which may be inadequate) or by fresh borrowing elsewhere (which will certainly be difficult) or by finding a buyer for the house (which may be even more difficult). The bank cannot be sure of immediate repayment—i.e., of converting the loan into money—and the loan, therefore, is illiquid.

**Differences in Interest-Rates.**

Here we have one of the reasons why different rates of interest are

charged on loans of different kinds. Loans are made for different periods of time. At one extreme, we have money deposited with a bank on current account; the money can be withdrawn without notice, so that the period for which it is lent to the bank is as short as depositors choose to make it. At the other extreme, we have bonds such as British Government Consols which the owners can never ask to have repaid; the holders of Consols have lent their money for an infinite period of time—although they can generally recover their money without difficulty whenever they need it by finding someone to take over the loan, i.e., to buy the bonds from them. Between these extremes there is a wide range of loans made for varying periods. The longer the period of the loan, the less liquid it will be, and the higher, generally speaking, will be the rate of interest payable. In normal circumstances, there will be a family of interest rates, increasing in magnitude from nil on bank deposits to a maximum on irredeemable bonds. A British Government loan maturing in forty years will normally require to be floated at a higher rate of interest than one maturing in twenty years and this in turn will carry a higher rate than Treasury Bills which mature in three months. In abnormal circumstances, however, rates of interest may be higher on short-term than on long-term loans. In a financial crisis, it may be impossible to borrow money on good security for three months at less than 6 per cent., while loans for three years are still being made at 4 per cent. This preference on the part of lenders for a less liquid and apparently less remunerative investment is puzzling. But it simply reflects a general expectation that when the crisis passes rates of interest will fall again, and that the fall will be heavier on short-term than on long-term rates. Suppose, for example, that the normal rates of interest are 2 per cent. and 3 per cent. on loans for three months and for three years respectively, and that lenders expect a return to these rates at the end of three months. Then it is easy to see that by choosing the less liquid investment during the crisis, lenders can obtain a return of 4 per cent. (compared with the normal 3 per cent.), whereas by choosing the more liquid investment they are able to average only 2¼ per cent. (compared with the normal 2 per cent.). A comparatively small change in long-term rates, therefore, is sufficient to compensate for a large but ephemeral change in short-term rates. All that is necessary, indeed, is that people should *believe* the change in short-run rates to be ephemeral; for this belief will rob high short-term rates of their attractiveness and leave the supply of long-term loans undisturbed at much the same price as before. Short-term rates, therefore, are much more variable than long-term rates. It requires a revolution in the outlook of lenders to produce a major change in long-term rates of interest; but short-term rates are extremely sensitive to the state of trade and of confidence.

A second reason for differences in interest rates is that all borrowers cannot offer equally good security. Where the risk of default is negligible—e.g., on British Government securities—loans will be made

at rock-bottom rates. But whenever lenders feel doubtful of the borrower's honesty, or of his financial strength, they will be reluctant to lend, and will either refuse altogether or will insist on a higher rate of interest to cover the risk of default. They will charge a premium in excess of the rate payable by borrowers whose credit is irreproachable. This premium will vary with the standing of the borrower, his past record, and the pledges which he can offer as guarantees of ultimate repayment. It will vary, too, with the period for which loans are made, so that long-term loans are not only less liquid but also less secure.

Differences in interest rates may also be due to market imperfection. The capital market is made up of a great many sub-markets specialising in different kinds of loans and not always in close competition with one another. The banks cater for one kind of borrower, the building societies for another. The insurance companies cater for one kind of lender, the investment trusts for another. The market for short-term loans is cut off from the market for long-term loans. Thus borrowers and lenders, attached by habit or by ignorance to one sub-market, may raise or make loans on terms less favourable than the rates ruling in other sub-markets. Lenders, for example, may continue to make long-term loans at low rates of interest without giving adequate thought to the alternatives offered in the market for short-term loans; or they may put their money in a savings bank as a matter of course without considering the terms offered by building societies for deposits or by local authorities for short-term loans.[1] They may lend to business associates at less than the competitive rate so as to build up goodwill, or to friends at a nominal rate as an obligement. In the same way, borrowers may continue to raise capital through some customary channel at more than the market rate of interest. They may lack knowledge or experience of alternative methods of finance. Or they may wish to give their customers an interest in the success of their business and offer them favourable rates for loans.

Market imperfection is greatest when the obstacles to competition are geographical. The rates payable in one country, or in one part of the country, are often higher than elsewhere because borrowers have access only to the local capital market, and have either no experience of borrowing in other markets or are not sufficiently well known to lenders outside their own district. Finally, differences in marketability lead to differences in the interest payable on large and on small loans. A company which makes a small issue of bonds may have to pay a comparatively high rate of interest because, since trading in these bonds will be restricted, there will be greater difficulty in disposing of them at short notice when lenders wish to recover their money and also because, in a narrow market, the value of the bonds will tend to fluctuate more abruptly.

[1] Between 1914 and 1918, for example, withdrawals from Savings Banks in Great Britain were comparatively trifling although much higher rates were obtainable on the loans issued during this period by the Government.

**Fixed and Free Capital: The Mobility of Capital.**

The savings of the past have already been invested; they are fixed in the stock of concrete capital which society has accumulated. The savings of the present are still free and undecided in their use; they are available for financing additions to the stock of capital in those industries and districts where the demand for capital is most urgent. If the demand for capital increases in any one use, therefore, it is mainly out of "free" capital that this demand is met. The rate of interest tends to be forced up, choking off demand elsewhere, and diverting savings to the expanding industry or district from other industries and districts in which these savings would normally have been invested. Just as the mobility of labour is greatest amongst new entrants, so the mobility of capital is greatest when capital is still free. But fixed capital, too, may be transferred from one use to another. If the production of one commodity is discontinued, the buildings and machinery used in its manufacture may be adapted for use in the production of some other commodity. The same plant may turn out a variety of products any one of which may be substituted for the others without dislocation; shop space may be used for the retailing of a fluctuating range of goods; a dismantled ship may supply a hotel with furniture and panelling, a coal-mine with machinery, and a steelworks with scrap. Thus one industry may decline and others take its place without drawing extensively on free capital. A third element of mobility arises through the conversion of fixed into free capital. As fixed capital wears out, a sum of money is set aside annually to provide for its ultimate replacement. These sums we may call "depreciation funds," since they are supposed to provide for current depreciation. Now depreciation funds are really free capital. They can, if their owners so choose, be reinvested in *other* kinds of property—in other industries and districts—which are yielding a high return, instead of being applied automatically to the replacement of property which may be yielding a comparatively low return. There is no *necessity* to replace property as it wears out and so preserve unchanged the physical stock of capital. Indeed, it is very unusual for the new property to be precisely similar to the old; when plant is replaced an improved model is generally introduced. But it *is* rather unusual for the owners of depreciation funds to withdraw them from a comparatively unprofitable business and invest them in some other business which offers a better return. From motives of optimism or of pride in his business, the company director—and still more, the individual *entrepreneur*—will continue to renew buildings and plant in spite of the attractiveness of competing opportunities of investment. Thus although in an industrialised community depreciation funds are considerably greater in amount than current savings, it is savings which make the greater contribution to capital mobility. It is not so much through the leakage of capital out of declining into expanding industries as through the diversion of savings into expanding industries that adaptation to changes in the demand for capital is secured.

**Why Interest is Paid :**

**(2) The Demand for Capital.**—Since borrowers are willing to pay interest, it is to be presumed that loans render some service at least equal in value to the interest paid. What is this service ? By what process does the £100 which I borrow this year become the £105 which I shall be asked to repay next year ? Do loans, in some mysterious way, increase the product of industry, or is the additional £5 paid out of the profits of exploitation by robbing workers of the full value of their labour ? Few questions in economics have given rise to so much controversy.

The most famous exponent of the exploitation argument was Karl Marx. According to Marx, the only productive agents are human labour and natural forces. Capital is not an independent factor of production; a loan by itself can obviously add nothing to the output of industry, and if the borrowed money is used to purchase capital equipment or raw materials, the equipment and materials are, in the last analysis, the product of man and nature. Machines, for example, are made by man from mineral deposits with the help of other machines; these machines in turn were made with the help of earlier machines. If we go back far enough to the first tools with which machines were made, these tools must have been the direct product of human labour. Machinery, therefore, is stored-up labour, and this past labour should be paid for at the same rate as present labour, without any addition of interest or profit. Capital is reducible to the labour and land which it embodies and renders no independent service. Nature, which partners labour in production, asks no compensation for its services. The full value of what is produced, therefore, is labour's by right; anything that goes in rent, interest or profit to the landowner, or to the capitalist, is seized from labour by exploitation.

This is a plausible argument, and one with a strong emotional appeal. But there is a flaw in it; it overlooks the value of time. Now time is precisely the service which is rendered to producers and to consumers by capital. We recognise this clearly enough when we are given credit—that is, when a loan frees us from the necessity of making immediate payment. For the privilege of paying at some more convenient time, we are willing to allow interest to be added to our bill; in other words, we are willing to buy time. But this is far from being the only kind of transaction in which time is of value. In nearly all productive processes an interval of time must elapse between the work which we do and the consumption of the finished product. There is an interval between ploughing and sowing, between sowing and reaping, between reaping and milling, between milling and baking. There is an interval between the commencement and completion of a house, a ship, or a machine, and a still longer interval before the full value of such durable goods is exhausted in the shelter or transport or manufactured goods which they yield over their years of service. Thus the efforts of a farmer or a builder bring in no immediate return; they fructify at some more or less

distant date. It is necessary to *wait* for that fructification, and this waiting is possible only if someone saves. The farmer must support himself and his workers until his crop is sold, and if he is to obtain command of the necessary purchasing power he must either save himself—i.e., postpone consumption—or borrow from some other saver—i.e., induce someone else to postpone consumption. To the work done by the farmer, therefore, there must be added the waiting done by the saver. Waiting is just as indispensable a constituent of production as working. Capital, which is the product of waiting, is as much a factor of production as labour, by which working is undertaken. To put the point another way: capital cannot be reduced entirely to past labour. For combined with the past labour which is stored up in machinery and other capital goods is the waiting that must be done before that labour repays itself.

The longer the interval of time between effort and return—what is sometimes called the "period of production"—the more effectively we can work. Suppose, for example, that Robinson Crusoe (whom no one can exploit) goes fishing. The most primitive method which he can use is to catch fish with his hands, as boys do. Then he may use a more roundabout method—carve a spear, or take a day off to make a rod and tackle and look for bait. Instead of spending all his time fishing, he now divides up his time between making fishing gear and using it. Later, he may use still more roundabout methods, building a boat and making nets for himself before resuming his fishing. Each change of method involves the use of an increasing amount of capital—first a spear, then a rod, then a boat and nets. Thus an increasing amount of capital is the product of an increasing amount of waiting; Crusoe's exertions in building the boat, for example, add nothing to his catch until the boat is ready for the water. But these exertions will *ultimately* increase his catch by far more than he could ever have caught by more primitive methods in the time which he gave to the work of construction. He is more than compensated for this work because of the technical superiority of roundabout or capitalistic methods of fishing. But he can only use such methods if he is able to wait for the greater reward that boat-building will ultimately bring him. If, in a less deserted island, he could borrow enough to pay for a boat, he would be able to transfer the burden of waiting to other shoulders by the offer of interest; and this interest would come out of the increased catch which he could make with the help of the boat.

This illustration can readily be generalised. Most commodities can be produced by a variety of methods, some requiring a great deal of capital, some comparatively little. Producers can use first very primitive methods with easily constructed tools; then more elaborate methods with simple types of machinery; then power machinery erected in expensive factory buildings. At each stage, output per man increases—not simply per man operating the machinery, but output per worker, including engineers, bricklayers, and all who build and repair the factory equipment. Although

output per head is greater, the product is not available for consumption until a longer period of time has elapsed. The work which is done (e.g., by engineers) is directed to the satisfaction of wants that are increasingly remote as the period of production is extended. There is more "jam to-morrow," but only if we are willing to wait until to-morrow. On the other hand, therefore, the use of capital increases the time-interval between taking the first steps in production (e.g., making tools) and turning out the finished goods, and so permits of the introduction of more efficient methods. On the other hand, since these methods, far from increasing current output, are made possible only by the withdrawal of resources from meeting immediate wants—as one withdraws eggs for incubation—they cannot be adopted without sacrifice—the sacrifice involved in waiting or saving. Employers are willing to pay interest because of the technical advantages of lengthy and roundabout methods of production; they are forced to pay interest because the burden of saving which the community is willing to carry is a limited one.

Capital is required, then, because production takes time; a given number of men can produce more than twice as much at the end of two years as they can produce at the end of one, provided they can maintain themselves out of savings for the extra year. Capital may be required also for other reasons. First of all, goods may grow in value over time, without any additional expenditure of effort. Given time, the forces of nature will cause wine to mature and trees to grow, unassisted by man. The function of capital here is not to allow of the introduction of a more roundabout process, but to allow us to wait while nature does its work; the technique is of nature's devising, not man's. Secondly, capital may enable goods to be transferred from a time of plenty to a time of scarcity, or from a time when they are little in demand to a time when the need for them is comparatively great. Potatoes are plentiful in September and scarce in April; the capitalist, therefore, by buying potatoes in September, and *waiting* until April, renders a service exactly similar to the services of traders in buying potatoes in *places* where they are abundant and carrying them to *places* where they are scarce. The greater scarcity of potatoes in April will raise their price and leave the capitalist with a profit, part of which is really interest on his capital.

Thirdly, capital may allow us to enter into immediate possession of goods which we urgently need. Not having sufficient capital ourselves, we borrow from other people and pay them a premium in interest for taking over from us the burden of waiting. The man who buys goods on the instalment system, the spendthrift who runs into debt, and the government which raises a war loan, share a common preference for present rather than future goods. They are disposed to dis-save, not to save; to anticipate their future income instead of trying to add to it. Each of them is using up savings, and none of them—apart, possibly, from the man who buys on the instalment system—is adding to the capital stock of the community

as a whole. Nevertheless the loan of capital is of as real service to them as loans made for other purposes, and the interest paid is in recognition of this service.

**The Marginal Productivity of Capital.**

Capital, then, does render a service to those who borrow it; it *is* productive. Its productivity falls, however, as the supply increases; capital, like the other factors of production, is subject to diminishing marginal productivity. Each fresh extension of the period of production, for example, yields a diminishing increment of product. When capital is scarce, only those time-consuming methods of production which promise great ultimate gain can be adopted; when more capital becomes available it is possible to use methods which repay themselves less handsomely. With a little capital we can generate electricity from coal and obtain all the advantages of electric over labour power; with more capital we can reduce the cost of electricity by the construction of dams and of hydro-electric power stations. With a little capital we can buy a second-hand motor-car; with twice as much capital we can buy either a new car or a higher-powered second-hand car. For the additional outlay we enjoy added comfort or speed. But we probably rate this added pleasure below the pleasure provided by our original investment; our enjoyment is not increased in proportion to the cost of the car. This tendency to diminishing productivity is almost universal. With more capital a merchant can keep a larger stock on hand, or a wider variety of goods; a manufacturer can use machines of better quality or more capitalistic techniques. But the additional capital yields a smaller return than the existing stock. The chief exception to this rule is the borrowing of governments for purposes of war; a larger supply of capital might make all the difference to such a government between success and failure, so that additional borrowing would be of greater service, not less.

The productivity of capital, in this context, means the productivity of *loan*-capital, i.e., the productivity of the things which can be bought out of a loan. The marginal product of capital is the additional return, after allowing for depreciation, which the borrower expects to obtain through the *use* of an additional unit of loan-capital. The marginal productivity of capital varies, therefore, not only with the annual return expected on capital equipment, but also with the cost of such equipment. If, as capital accumulates, machinery becomes more costly in relation to the price of its product, the marginal productivity of capital will fall, both because the uses to which the additional machinery can be put will be less urgent, and because a greater investment of capital will be necessary for the purchase of the machinery.

In many uses to which capital is put its productivity is not easily measured. Who can say, for example, whether houses are more productive than roads, or roads than battleships? The measuring-rod of money can be applied to houses to give the annual return on the

capital invested, but roads and battleships yield no monetary return. Are we to say, therefore, that houses alone are productive? Clearly, this would be absurd. What we must say is that productivity is measured by the borrower himself—sometimes, when he is producing goods for sale, with an eye to the judgments of productivity expressed by consumers in the prices which they pay; and sometimes, when he is meeting his own immediate wants, on the basis of his private judgment of how these wants can best be met. A business man's judgment of productivity rests on his expectations of profit—which is measurable; a government's judgment rests on its conception of the needs of the community—which may not be measurable.

The marginal productivity of capital governs the rate of interest which borrowers are prepared to pay. Producers will put capital to uses in which the prospective return becomes progressively smaller as more and more capital is applied. Under perfect competition, they will increase their borrowings until, on a further loan, the prospective return over the period of the loan is a trifle less than the interest payable on it—that is, until the marginal productivity of capital and the rate of interest are equal. The return which producers balance against the interest cost of a loan is generally obtained through the conversion of the loan into concrete capital assets. Since the contribution made by these assets to the value of output extends into the future over the period of their life, the gross return is not certain, but must be estimated by producers on the basis of their experience of the past and their anticipations of the future; it is a *prospective* return. Since, secondly, capital assets depreciate over their life, provision must be made out of the gross return for the probable cost of depreciation; it is the *net* return which is balanced against the interest charge. Thirdly, the return will vary according to the period for which the loan is made. Producers can generally turn long-term loans to better account than short-term loans, and will generally be willing, therefore, to offer comparatively high rates of interest for long-term loans. For some purposes, however, they will prefer to borrow for short periods even at high rates; for instance, to finance a temporary, or a seasonal, increase in the stocks of raw materials or of finished goods which they wish to carry, or to tide over a trade depression, or to provide themselves with liquid funds for emergency payments. The rate of interest on loans for any given period of time will tend to be equal to the marginal productivity of capital invested for that period of time.

When competition is not perfect, the marginal productivity of capital will exceed the rate of interest for two reasons. If there is imperfect competition in the market for loans, borrowers may find it impossible to obtain command of additional capital except by paying successively higher rates as their borrowings increase. They will tend, therefore, to restrict their borrowings since each additional loan, while adding to the value of output, not only adds to their costs the interest charge on the new loan, but also increases the interest

charge on previous loans.[1] If there is imperfect competition in the market for goods, borrowers, as we have seen, will restrict output so as to keep up prices, and in restricting output will also restrict their borrowings. They will set against the value of the goods produced from additional capital, the fall in the value of the goods produced by the existing stock of capital. Thus the *net* return to a marginal application of capital will fall below the marginal productivity of capital because the consequent increase in sales will drive down prices.

**Changes in Marginal Productivity.**

The demand for capital will be increased whenever there is a rise in the marginal productivity of capital. Such a rise may occur as a result of invention; for example, the invention of new means of transport (the railway and the steamship in the nineteenth century, the motor-car and the aeroplane in the twentieth), or of communications (by cable, telephone or wireless); or of new sources of power (steam, electricity, petrol and oil). Secondly, the marginal productivity of capital will be increased by anything which makes concrete capital scarce relatively to the other factors of production—the destruction of property on a large scale in an earthquake or a war; an increase in population; an increase in the efficiency of labour; the discovery of new natural resources—minerals, oils, etc.—and the settlement of new and undeveloped countries. Thirdly, the marginal productivity of capital will be increased by anything which disposes people to anticipate their income and consume the savings of others; governments, for example, may borrow heavily to finance preparations for war and the conduct of the war itself. Fourthly, the marginal productivity of capital will be increased by a change in tastes from goods which require a low proportion to goods which require a high proportion of capital in their manufacture: for example, people may spend more on rent and move to larger houses, which require the investment of a great deal of capital; and spend correspondingly less on drink into which capital does not enter so largely. *Any* change in demand, indeed, will be likely to raise the marginal productivity of capital. For it will be necessary to provide more of one commodity and less of another; and since the machinery and plant used in the manufacture of the second commodity is unlikely to be adaptable for the manufacture of the first, more machinery and plant will have to be constructed if the output of the first commodity is to be increased. The demand for capital, therefore, will increase and the increase will be all the greater if a comparatively high proportion of capital to labour is used in the manufacture of the first commodity. Fifthly, the marginal productivity of capital will be increased by anything inducing

[1] This assumes that they pay the same rate of interest on all their borrowings. If they can discriminate between one lender and another, they will continue to borrow until the rate of interest *on the last loan* made to them is equal to the marginal productivity of capital.

business men to take a more optimistic view of the future. If, for example, producers are in a more optimistic frame of mind, they will entertain more favourable expectations of profit; this will raise the prospective return on capital and stimulate borrowing. An important influence making for optimism is past experience of a steadily expanding market. The marginal productivity of capital is higher in a world of expanding population and trade than in a world where population is tending to become stationary and trade restrictions accumulate. It is higher not only because business men tend to be more adventurous, but also because, whatever their mood, they can justifiably look forward to an expanding return from year to year. Thus we can add one last type of change in marginal productivity to our list—dynamic changes deflecting demand from a given trend. These changes may be associated with the growth of population, or of trade barriers; they are chiefly important in the short period, when they are associated with fluctuations in employment and in income.

**What governs the Rate of Interest?**

So far we have made no attempt to formulate with any precision the forces controlling interest rates. At one time interest rates have been made to depend upon the state of liquidity-preference of the public; their *rôle* has been to equalise the advantages of holding money and loans. At another time, attention has been focussed on the supply of loanable funds and the demand for their use, and the *rôle* of interest rates has been to preserve equality between the two. The first explanation is in terms of the existing *stock* of loans, the second in terms of the *flow* of loanable funds. But what appear to be two distinct and conflicting explanations are really different versions of the same explanation.

According to the liquidity-preference version the public must choose between holding its wealth in the form either of loans or of money, or of other assets.[1] Given the supply of money and of loans, the demand for each must be such as to absorb this supply. The price of loans in terms of money, therefore, will vary until the public has no incentive to change from loans to money or from money to loans. But the price of loans is simply another expression for the rate of interest. The rate of interest, therefore, is governed by the demand for money relatively to loans (liquidity-preference) and by the supply of money relatively to loans. The rate of interest will change for one of three reasons—a change in liquidity-preference, a change in the supply of money, and a change in the supply of loans. Now of these three reasons the last is equivalent to a change in the demand for loanable funds while the first two are equivalent to a change in the supply of loanable funds. Suppose, for example, that there is an increase in the marginal productivity of capital. Then more will be borrowed and the rate of interest will rise—from one point of view, because the consequent addition to the existing stock of loans

[1] Other assets are ignored here for the sake of simplicity.

forces down their price in terms of money; from the other point of view, because increased borrowing puts greater pressure on the public's willingness to lend and hence on the supply of loanable funds. Or suppose that there is an increase in the supply of money. The increase, which may be made through inflation of the currency by the government or through the grant of additional credit by the banks, will add to the supply of loanable funds and reduce their price; alternatively, it will add to the stock of money, and diminish the price at which people are willing to purchase liquidity. From either point of view, the increase in the supply of money will operate to depress interest rates. Finally, a change in liquidity-preference clearly implies a change in the public's willingness to lend. If people are less anxious to hold their wealth in monetary form they will be more willing to lend it. The supply of loanable funds, therefore, will increase whenever there is a reduction in liquidity-preference.

It will be observed that in the list of forces controlling interest rates saving is not included. Yet we know that without saving there would be no capital to lend. A theory of interest rates that ignores saving is like Hamlet without the Prince of Denmark. Why then has saving been left out? The reason is that saving influences interest rates only through other factors. If more is saved, for example, the supply of loanable funds may increase. On the other hand, the demand for loanable funds is likely to decrease. If people become more thrifty and eat less bread or stay in smaller houses there is less need of ovens and dwelling-houses. Thus the incentive to borrow for the erection of ovens and dwelling-houses is reduced, and with it the demand for loanable funds.

Although there is a close connection between the amount which the community succeeds in saving and the supply of loanable funds, this does not mean that increased thrift necessarily increases the supply of loanable funds. For increased thrift does not necessarily or automatically increase the amount saved by a community, although it does necessarily increase the amount saved by any one person. It is within the power of one man to add to his savings by increased thrift. But it is not within his power to add to the total savings of the community. His decision to save more may cause someone else to save less. If, for example, he spends less on bread, the baker will make a lower profit and will have less income out of which to save. If the baker tries to save as much as before, he in turn will spend less and reduce the profits of some other tradesman. And so it may go on, with each man trying to save as much as before and in doing so reducing his neighbour's income. An excess of thrift does not of itself increase the amount saved or the supply of loanable funds; it may, instead, make for industrial depression and unemployment. Only when additional thrift is compensated by additional loan-expenditure is there a net increase in savings. For then the income of the community is undisturbed; some people lose because the savers spend less, while others gain as the borrowers spend more.

The fact is that the amount which the community saves does not depend upon conscious decisions taken in advance by the savers. For savers cannot, in advance, decide what their incomes will be; and savings, as we have seen, depend upon income. We may plan what we mean to save, but since we do not know what others are planning or what the future holds for us, our realised and our anticipated incomes may differ; some part of what we save will be undesigned. On the other hand, nothing of what we *lend* is undesigned; we can plan at any moment how much of our wealth we will lend and how much we will retain in the form of money and other assets. Thus we can draw a supply curve of loanable funds, showing how much we will lend at any given rate of interest. We can also draw a supply curve of planned savings to show how much we will plan to save at any given rate of interest. But we cannot draw a supply curve of amounts actually *saved*, for a supply curve is the expression of an attitude towards the future, not a statement about the past.

Here, for the time being, with many puzzling questions unresolved, we must leave the theory of interest. Later, when we come to discuss industrial fluctuations, we shall return to the subject. We shall see that the rate of interest, in addition to being the reward of loan-capital, is also one of the most important regulators of our economic life and that upon our control over interest rates depends in no small measure our ability to make full and continuous use of our productive resources.

## 2. Profit.

Profit, as generally understood, is the difference between the total expenses incurred in producing or acquiring a commodity and the total revenue accruing from its sale. This difference may be expressed as a return on *capital*, the total profit over a year being related to the amount of capital employed; alternatively, profit may be expressed as the proportion by which the price per unit sold exceeds its cost, i.e., as a rate on turnover. A village chemist, for example, may be able to make a very large profit on turnover and only a moderate return on his capital; whereas a firm of wholesalers may make a comparatively trifling margin of profit on turnover, and yet obtain an abnormally high return on its capital. Profit on turnover is of no significance in the theory of distribution. It is profit *as a source* of *income* that we are studying, and since income is a flow measurable over time, profit can form part of income only if it, too, is measured over time instead of being expressed as a ratio on individual transactions. We shall confine ourselves, therefore, to profit as a return on capital.[1]

[1] Profit on turnover can be derived from profit on capital by dividing by the rate of turnover of capital. If, for example, capital is turned over twice annually and annual profits are 6 per cent., then the rate of profit on turnover is obviously 3 per cent.

**Gross and Net Profit.**

Profit as defined above is described by economists as *gross* profit, because it may include some items which are not really profit at all. The first of these is earnings of management. A joint stock company includes in its expenses (and so excludes from profits) the salaries of those who undertake the work of management. But in businesses which are managed by their owner (e.g., farms, shops, etc.) the same kind of work is done without expense, and profits are swollen therefore, by the value of this work. The true or *net* profit can be arrived at only after deduction of the expense to which the owners of such businesses would be put if they had to hire the services of a manager; or, alternatively, after deduction of the salary which the owners might obtain by doing equal work for a joint stock company.[1] Secondly, gross profit may include what ought properly to appear under the heading of interest. A business which uses borrowed capital counts the interest on this capital as an expense and excludes it from profit. Other businesses, using their own capital, include in profit the whole return on their capital. The proper procedure, if we wish to isolate net profit, is to deduct from the gross return on capital the interest which might have been earned by lending that capital on good security. A third item which ought properly to be excluded is the rent of land or buildings owned by the firm. A farmer, for example, might fail to reckon the rentable value of his land as a business expense in arriving at his profit on the year's working.[2]

**Distinguishing Features of Profit.**

Profit differs from other kinds of income in three ways. First, it may be negative. Neither wages, rent nor interest are ever likely to be negative, but every year there are some firms which make a loss and there are few firms which do not make a loss at some time or another. Some economists even take the view that, taking into account business failures, net profits are, on the average,

---

[1] Strictly speaking, a rather larger deduction should be made. It costs something to bring capital and business ability together, so that an important advantage is lost by any business man who lends out his capital and hires himself to a joint stock company for a salary. This advantage, which is similar to the advantage of favourably situated land, gives rise to an element of rent which should be included in earnings of management rather than in net profit.

[2] The distinction between gross and net profit is by no means an academic one. When, for example, a tax is levied on profits, the usual basis of assessment is gross profit. This obviously discriminates against businesses managed by their owners or employing their owners' capital. There is also, as we shall see, discrimination against young, enterprising and growing firms in comparison with old and established ones. A tax on *excess* profit, —i.e., on profit in excess of some "fair" or "normal" return—may also involve discrimination against firms which are making a high *gross* profit and only a normal *net* profit.

negative—that is, that those who expose their capital to the uncertainties of business could earn a higher return, on the average, by lending out their money at interest. The statistical evidence, which is not very adequate, suggests that this view is too extravagant. But there can be no doubt that the *average* net profit which is earned in business (after deducting all losses) is comparatively small, and that it is greatly exaggerated by those who think too exclusively of a few exceptional gains. Secondly, profit fluctuates more than any other kind of income. Between boom and slump there is comparatively little change in wage-rates (or even in wage-incomes), in interest-rates (the change in long-term rates is generally trifling), or in rents; the brunt of the change falls on profit. Profit responds immediately to a change in price; other incomes are adjusted more slowly and less violently. Thirdly—and this is the crucial distinction—profit is not, like other kinds of income, a contractual and certain income, agreed on in advance, but an uncertain residue determined by the course of events. A man's wage, for example, is predetermined and certain in amount; but the income of his employer is not. The man is paid now for goods produced in anticipation of future demand; and, since the future can never be foreseen with certainty, the goods which the employer obtains for a given wage-payment are of uncertain value. If eventually the price which the goods realise exceeds the cost incurred in their production, the employer will make a profit; if the selling-price turns out to be lower than the cost which he has incurred, he will make a loss. But he does not, and cannot, know in advance that he will be able to make a profit. He *expects* to make a profit, but his expectations may prove false.

Profit is the difference between two prices, one of which is present and certain, while the other is future and uncertain. This is most obvious of mercantile profit—the profit of traders. Traders acquire goods at one price in the expectation of selling again at a higher price; the price at which the goods can be bought is known and certain, but the price at which they can ultimately be disposed of is uncertain and depends upon market risks which each trader must judge for himself. Similarly a financial profit is made by the purchase of property or securities at one price and selling again at another price which is expected to be higher but which cannot be predicted with certainty. Industrial profit arises in the same way—either current outgoings are known and future revenue is incalculable, or, when work is done under contract, the contract price is known and the cost is not.

**The Origin of Profit.**

Thus profit originates in uncertainty. That uncertainty arises out of the responsibilities of ownership in a world of change and imperfect foresight. If everything could be reduced to routine, or if the future could be exactly predicted, there would be no uncertainty, no need to take chances, and no profit in the strict sense.[1] But in the world

[1] There would, however, be *monopoly* profits. As to these, see pages 186-7.

as we know it, there *is* uncertainty, the burden of which is all the greater because of our intricate economic system with its intense specialisation and its production for markets distant in time and space. The burden of uncertainty is one which *someone* must bear. Part of the burden—the only part under discussion—is borne by the owners of property.[1] But it requires the lure of profits to induce them to bear it; they will not run the risk of loss except in the hope of eventual gain. There may be some people with such a liking for uncertainty that even the prospect of loss would not deter them from indulging their adventurousness; but most people take chances less from love of gambling for its own sake than from hope of making a profit if they are lucky. It does not follow from this that profit must be on the average a positive quantity. But it does follow that some people must make a profit and others a loss; there must be a prospect of profit which more than compensates for the risk of loss.

**Profit and Cost.**

People will take chances, therefore, only if they judge the chances to be in their favour—that is, only if they expect to make a profit. Before they will invest their capital in any industry they must have a sufficiently strong expectation of profit to overcome their fear of loss. This prospective profit is a cost exactly like wages, interest and rent. It is a necessary inducement without which adequate supplies of capital would not be obtainable. Prospective profit, therefore, enters into the supply-price of finished goods; the price anticipated must be high enough to cover wages and interest charges and leave a margin which is considered adequate by business men to cover the various contingencies to which they are exposed. This margin will be larger for some business men than for others. Those who are temperamentally cautious and averse to the assumption of risk will work to a larger margin than those who are more confident and venturesome. Those who take a gloomy view of the probable trend of prices will hesitate to take risks which seem moderate to more optimistic competitors. Thus there are marginal risk-takers, just as there are marginal savers and marginal workers; and it is the marginal risk-takers to whom prices must offer an adequate prospect of profit.

Although prospective profit enters into cost, realised profit does not. It is the reward for successful risk-taking and is a surplus governed by, not governing, price. At the same time, since men's expectations are

[1] The burden of uncertainty is not borne exclusively by property-owners. An important part is borne by labour, since workers run the risk of injury, ill-health, or unemployment. The burden on labour, however, has been relieved by schemes of social insurance and by State expenditure on the social services. Part of the burden on capital has also been transferred to the State, since the grant of subsidies to depressed industries relieves the capitalist of the fear of financial loss. In addition, the State assumes the burden of uncertainty for all property in its ownership: so that the burden is at a maximum in a socialist community and increases with every extension of public control and operation of industry.

generally grounded on their experience, prospective profit and realised profit are likely to move together. In the long run, if an industry consistently fails to yield a profit, business men will give up expecting it to yield one. By influencing expectations, therefore, and so controlling prospective profit, realised profit in the long run enters indirectly into price.

Realised profit is partly the result of luck, partly the result of good judgment—since some people take foolish risks—and partly the result of the skilful avoidance or elimination of unnecessary risks—i.e., of outstanding organising ability. But if anyone shows evidence of good judgment and is right oftener than other people, the return which he obtains on his capital is not all pure profit but a combination of profit and earnings of management. For a man of sound judgment can hire his services for a salary to people who have no faith in their own judgment or who recognise their own incapacity. In the same way if one man is more skilful than his fellows in *avoiding* or *eliminating* risks because of his superior powers of organisation he will earn a higher profit which reflects, not better luck, but greater talent. Once the return to superior organising ability, like the return to superior judgment, has been transferred to earnings of management we are left with a residue of pure profit which is the net return to uncertainty-bearing and is governed solely by luck.

**The General Level of Profit.**

When there is great uncertainty the rate of profit tends to be high. In forecasting the future, people will leave themselves a wide margin of error, and in normal circumstances this will be reflected in a correspondingly high margin of realised profit *on the average*. For *individual* risk-takers, however, profit will range between a very large positive and a very large negative quantity; in a new industry, for example, some firms generally earn enormous profits while others lose their capital altogether. When uncertainty is comparatively negligible, the average margin of prospective, and of realised, profit will be correspondingly small and profit will range within fairly narrow limits; businesses supplying some staple article of clothing or of diet, for example, can generally predict their future profit with much more assurance than firms of shipbuilders or steelmakers. The correspondence between degree of uncertainty and prospective profit—and *a fortiori* between uncertainty and realised profit—is by no means perfect. Many people are more attracted by a spectacular chance of sudden fortune than by humdrum risks which are not likely to cause heavy loss and are equally unlikely to bring great gain; some outsiders, so to speak, are always more heavily backed than they should be. Thus profit is often highest in trades where risks are moderate—neither so negligible as to make risk-bearing almost superfluous, nor so enormous as to attract the incautious and the speculative. It was neither the railways nor the gold-mining industry which earned the highest profits in the late nineteenth century; it is neither the omnibus companies nor the film industry nowadays.

**What Keeps up Profit.**

The degree of uncertainty, and hence the rate of profit, will be higher the greater the rate of social chance and the greater the degree of immobility of productive resources. The more rapidly change takes place, the more difficult and dangerous is the task of predicting the future and of undertaking the adjustment of resources to future wants. Whether that adjustment is planned by the State—as under socialism—or undertaken by private enterprise—as under capitalism—a larger margin of profit will require to be added to present costs to provide for the uncertainties of the future. The task of adjusting resources to wants will be greatly increased if these resources are specialised and immobile; for once they are committed to one use it will be difficult to adapt them for other purposes when new information or discoveries call for a change of plans. If resources are immobile the burden of uncertainty in deciding their use is heavy, and any acceleration of social change, by dictating more frequent and more drastic readjustment of resources to wants, will aggravate the burden. The risk of loss through sinking capital in forms no longer required is increased, and to overcome this increased risk of loss there will require to be a correspondingly higher expectation of profit.

**Profit and Progress.**

The rate of profit, therefore, is intimately connected with what is generally called "progress." It is the business of the *entrepreneur* to make progress by pioneering new methods and new products. If his innovations are successful he reaps a profit; if he is unfortunate or injudicious he loses his capital. Because progress means change, it creates uncertainty; and because progress means increased efficiency it creates a prospect of profit. Progress, therefore, is one of the chief forces maintaining the rate of profit and is itself one of the chief sources out of which profit is paid. There are, however, some ways in which progress makes for a lower rather than a higher rate of profit.

**What Reduces Profit.**

First of all, methods may be found of reducing specific risks. Wireless communication and weather forecasting have reduced the risk of shipwreck; agricultural research has reduced the risk of crop failure through destruction by pests or unfavourable weather; a granary reduces the risk of famine. Secondly, uncertainty may be reduced by devices which make specific risks more measurable. The more measurable risks are, the more certain is their incidence and the smaller is the burden of uncertainty that must be shouldered. If, for example, a manufacturer of golf-balls knows that one in ten must be rejected as below standard, then the certain cost of these discarded balls is simply added to their price; if he knows only that between one in eight and one in twelve must be discarded, then the loss ceases to be a known item in expenses of production and the addition to price becomes swollen by an uncertainty-premium;

if even the limits to the loss from defective balls are unknown, this greater incalculability will raise the uncertainty-premium still further. As risks become less measurable, the tendency is for a large uncertain loss to be substituted for a smaller more-certain loss, and for a larger payment for uncertainty-bearing to take the place of a smaller business expense. As risks become more measurable, payments to cover these risks approximate more and more closely to ordinary business expenses until a completely measurable risk gives rise to a known expense, free from uncertainty altogether.

Risks can be made more measurable by grouping. As we have seen,[1] it is often much easier to predict what will happen *on the average* than in some individual instance. We can predict fairly accurately how many houses will be destroyed by fire next year, but we cannot predict whose these houses will be. Anyone who owns a large number of houses, therefore, runs a more measurable risk of loss than the owner of a single house. The more houses one man owns, the larger is the number of similar risks which he groups together, and the smaller is the burden of uncertainty which he shoulders. The same result can be attained more easily by insurance. By the payment of a known and fixed premium to an insurance company, the householder can transfer the burden of uncertainty, while the insurance company, by grouping risks until they lose their incalculability, can estimate its losses fairly accurately, and can charge premiums which, while attractive to the insured, leave a margin of profit for the insurer.

Risks may also be made more measurable by bringing fuller information to bear on them. Adequate statistical data, for example, are essential in insurance. They are also essential in business. No one can form an accurate judgment on future trends without consulting the latest information on prices, output, unemployment, and so on. Since it is often difficult for one man to find time to collect and assimilate all the necessary information, specialised agencies come into existence to assist investors and business men in predicting the future; for example, trade journals, information bureaux, and consultants of various kinds. Where risks cannot be easily assessed without a mass of information which few people can take the trouble to acquire, another type of specialisation also grows up. Risks are assumed by specialists who from long experience acquire both skill in risk-taking and the special knowledge which is indispensable to successful risk-taking. These specialists are of particular importance in the markets for staple commodities such as cotton and wheat, where there are organised dealings in futures and hedging contracts; their function is to reduce uncertainty by more accurate assessment of risks.

Specialisation in risk-bearing is carried a stage further in the joint-stock company. Here there is a reduction of uncertainty through the delegation by the shareholders to the board of directors

[1] Above, page 75.

of the work of management, since the directors are likely to be better judges of market conditions and can take more trouble in obtaining information. There is also an increase in willingness to bear uncertainty because of the issue of different classes of share, catering to the individual propensities of different shareholders. This increased willingness to bear uncertainty is reinforced by statutory limitation of the shareholder's liability, which limits the risk of loss and therefore increases the probability of profit,[1]

**Who Receives Profit?**

The joint-stock company raises an important problem—who is it nowadays who bears the burden of uncertainty? In the nineteenth century one could point to the *entrepreneur*, who supplied capital, management and enterprise, all at once. In the twentieth century, however, each of these may be supplied separately—capital by debenture-holders, management by the board of directors, and enterprise by the ordinary shareholders. It might seem that uncertainty is now borne by the directors, with whom rests the responsibility for taking the day-to-day decisions to which risk attaches. But the directors are hired officials who do not bear the *ultimate* responsibility for their decisions. That responsibility rests with the owners—that is, with the preference and ordinary shareholders. They may delegate power to take decisions, and may even have no effective power to intervene, but it is they who none the less run the risk of loss and who enjoy any profit that is made.

**Variations in Profit between Industries.**

The tendency of competition is to make profits highest in those trades in which the burden of uncertainty is felt to be heaviest. Profits are likely to be high in industries in which methods of production are constantly changing, so that there is need for continuous adaptation of technique—in industries supplying luxury products the demand for which fluctuates rapidly; in industries which are still too young for their prospects to be judged with assurance; in industries in which a long interval elapses between investment and return of capital; in industries in which the labour and capital employed must be more or less irrevocably committed to narrowly specialised tasks; and in industries exposed to specific risks the incidence of which is uncertain—agriculture, for example, with its dependence on weather conditions, or brewing, with the threat of prohibition or of higher duties hanging over it. There is no tendency towards equality of profits, but only towards such rates of profit as equalise differences in the degree of uncertainty felt by investors in different industries.

**Monopoly Profit.**

Just as an element of rent may be included in wages or in interest, so profit may be maintained above the level which is necessary to

[1] For a fuller discussion, see above, page 90 *et seq.*

induce capitalists to bear uncertainty. This surplus profit has already been discussed under the heading of " producer's surplus." It arises in trades into which entry is not free and in businesses which enjoy some monopolistic advantage over their rivals. The general level of profit may also be maintained at an artificially high level if there are any restrictions on new businesses in general, or any unnecessary impediments to the acquisition of command over capital by enterprising men.

Profit, therefore, is divisible into two parts, one of which is to be explained along the lines of this chapter as a reward for uncertainty-bearing, while the other is to be explained along the lines of Chapter 14 as a monopoly gain.

# PART V.—INTERNATIONAL TRADE

## CHAPTER 22

## THE THEORY OF INTERNATIONAL TRADE

### TRADE AND SPECIALISATION

THERE is no fundamental cleavage between one kind of trade and another. All trade, whether between persons, towns, regions, or countries, originates in specialisation, and, whatever form specialisation assumes, the principles governing it are the same. A person, by specialising and exchanging, supplements his natural or acquired gifts out of the comparative proficiencies of his neighbours; in the same way, a region, by trading with other regions, supplements the resources with which it is most poorly supplied out of the comparative abundance of these resources elsewhere. The individual specialist, in order to make the best use of his limited time and energy, concentrates on a narrow range of tasks, relinquishing to others work which he could readily undertake himself; similarly, a region, in order to make the best use of its limited productive resources, specialises in the industries in which its comparative advantages are greatest and sacrifices other industries, in whole or in part, to competition from outside. It makes little difference whether the boundaries of the region are those of a county or of a country; the broad principles are the same whatever the frontiers across which trade flows.

**International and Domestic Trade.**

Regarded as the outcome of geographical specialisation, trade between countries is of no more interest than trade between other areas. Whatever the unit of area, much the same problems present themselves and much the same principles can be applied. It makes little difference whether we ask of a country or of a county why it does not pursue a policy of self-sufficiency, why it carries on this industry rather than that, and why it is a prosperous area when others are not. Indeed, there is a positive advantage in framing such questions in terms of local rather than of national units. For nothing makes so strongly against clear thinking about trade as the belief that international and domestic trade are fundamentally different from one another, and that, while it is natural and desirable to interfere with the one, it is equally unnatural and undesirable to interfere with the other. If we were to make a habit of recasting every argument about international trade in terms of inter-county trade we should be imposed upon far less readily by the gross and mischievous fallacies which colour popular economics. We should see the inconsistency of the American who supports tariffs on imports from low-wage countries, such as Japan, but never thinks of duties

on imports from low-wage States within the Union, such as Tennessee. We should see that an unfavourable balance of payments between South Wales or the Highlands and the rest of Britain has much the same significance as an unfavourable balance between Britain and other countries. And we should realise that, from the economic point of view, political boundaries are quite arbitrary; that much the same principles apply to trade between Berlin and Vienna when they are in different countries as when they are in the same country; that there was no major change in Ireland's trade when the north was severed from the south, nor in Britain's when England and Scotland were joined.

**Do we need a Separate Theory of International Trade?**

In spite of the fundamental resemblance between international and domestic trade there are also points of difference. But for these points of difference a separate theory of international trade would be unnecessary, since it would be nothing more than a simple application of the general theory of localisation. It would be sufficient to discuss (as we have already discussed in Chapter 5) what industries tend to settle in an area with given productive resources and on what scale these industries will be carried on. But countries are more than geographical areas; they are areas with economic systems of their own, cut off by political frontiers from the economic systems of their neighbours. They have, as a rule, their own currency and their own banking system. Their commodity markets are hedged about with customs duties and customs officers. Their capital-market is partially insulated from other capital-markets by restrictions on the export of capital and by the reluctance of investors to entrust their capital to distant enterprises and foreign governments. Their labour market is protected against immigration by visa restrictions and quota regulations as well as by the ordinary deterrents to long-distance migration such as ignorance of opportunities, travel expenses and language difficulties. Above all, each country has its own government, controlling trade on a national scale from the standpoint of national interests. We need a separate theory of international trade, therefore, both to take account of the comparative immobility of productive resources between one country and another, and in order to furnish general principles of national policy as a guide to action. First, we require to adapt the general theory of trade to the facts of immobility; and, second, we require to re-state it in terms of national units—the units of control.

**The Implications of Immobility.**

Since the factors of production cannot move freely between one country and another, they are often forced to remain in countries where their productivity is comparatively low, and where, in consequence, they are comparatively poorly paid. Workers who could obtain high wages if admitted to the United States or to Great Britain continue to accept much lower wages in Italy and Japan. Capital

which, if invested in India or Poland, might show a high return, continues to be lent at much lower rates to American and British enterprises. Moreover, since resources are not used where their productivity is highest, goods are not produced where costs are lowest. They are certainly produced where *money* costs are lowest, but not where costs are lowest in terms of real resources or of the efforts and sacrifices which go to production. It is far cheaper, in terms of human effort, to grow wheat in Canada than in India. But Indian farmers continue to grow wheat in competition with Canadian farmers. Whatever they grow, their efforts are less amply rewarded than they would be in Canada. So they choose crops which, although not better suited to India than to other countries, are better suited to it than other crops. They choose crops in which, although they may be at an absolute disadvantage in comparison with other countries, they are nevertheless at a *comparative* advantage relatively to other crops. The real cost of growing such crops is higher than elsewhere, but the money cost is lower because a given amount of effort on the part of an Indian worker earns less than the same amount of effort by foreign workers of equal efficiency.

**Absolute and Comparative Advantages.**

To illustrate this distinction more fully, let us suppose that there are only two countries, Anywhere and Nowhere, that each produces only two commodities, sugar and salt, and that the only factor of production in either country is labour (land being free and of uniform quality). Let us also suppose that Anywhere is better suited to the production of sugar and Nowhere better suited to the production of salt. For example, let :—

| | | | | | | |
|---|---|---|---|---|---|---|
| 10 | days' | labour in | Anywhere | produce | 100 | units of sugar. |
| 10 | ,, | ,, | ,, | ,, | 80 | ,, salt. |
| 10 | ,, | ,, | Nowhere | ,, | 50 | ,, sugar. |
| 10 | ,, | ,, | ,, | ,, | 200 | ,, salt. |

Before trade begins, sugar in Anywhere will sell at four-fifths of the price of salt, while in Nowhere it will fetch four times the price of salt. There will be an obvious advantage to both countries, therefore, if Anywhere exports sugar in exchange for salt at any price within these limits. What price will actually be established will depend upon the strength of the demand for sugar and salt in the two countries. Let us suppose that, the cost of transport being negligible, the price in both countries settles down after the opening of trade at two units of salt per unit of sugar. Then Anywhere, by exporting 100 units of sugar will obtain in return 200 units of salt ; so that, by 10 days' labour in sugar production, it saves 25 days' labour in the production of salt. Similarly, by exporting 200 units of salt, which it took ten days to produce, Nowhere obtains 100 units of sugar which it would have taken 20 days to produce ; 20 days' labour in the two countries produce as much salt and sugar as formerly required 45 days' labour.

In this example, Anywhere has an absolute and also a comparative advantage in the production of sugar, while Nowhere has an absolute and also a comparative advantage in the production of salt. But suppose that 10 days' labour in Nowhere produce, not 200, but only 50 units of salt. Then Anywhere will have an absolute advantage in the production both of sugar and of salt. It will still be worth its while, however, to trade with Nowhere, for Nowhere retains its comparative advantage in the production of salt; it can produce as many units of salt as of sugar per day's labour, whereas Anywhere can produce only four-fifths as much. At any price between four-fifths and one, both countries will gain from trade. Suppose, for example, that the terms of trade fixed by the joint demand of both countries for sugar and salt are 10 units of sugar for nine units of salt. By expending 10 days' labour in the production of sugar and exporting it in exchange for salt, Anywhere can obtain in return 90 units of salt which would otherwise have taken 11¼ days to produce. Similarly, instead of spending 10 days' labour in producing 50 units of sugar, Nowhere can obtain just as much by trade through nine days' labour in salt production. The only real difference from the first example—apart from the smaller gain from trade—is that the lower productivity of labour in both industries in Nowhere reduces wages there below the level in Anywhere. In our first example, wages must be exactly equal in both countries since 100 units of sugar (produced by 10 days' labour in A) are assumed to be equal in value to 200 units of salt (produced by 10 days' labour in B). Given terms of trade more favourable to sugar, however, wages will be higher in Anywhere, and given terms more favourable to salt, they will be higher in Nowhere. In the second example, it is easy to calculate that wages in Anywhere will be nine-fifth times wages in Nowhere. Here again, however, relative wage-rates depend upon the terms of trade between sugar and salt, and may lie anywhere between ten-fifths and eight-fifths.

**Incomplete Specialisation.**

There is no suggestion in this argument that a country gains only through the import of goods which it cannot produce for itself. We have explicitly assumed that sugar and salt can be produced both in Anywhere and in Nowhere, and have tried to show that each gains from a policy of specialisation. The logical conclusion of our argument would, in fact, be complete specialisation, Anywhere giving up salt production and Nowhere sugar production. But specialisation is usually limited by forces of which we have taken no account. Both industries, for example, may be subject to increasing marginal costs, so that it is not worth while to expand sugar production in Anywhere beyond a certain point, while it is worth while to continue salt production on a modest scale. To push sugar production further would raise costs excessively, while the cost of producing salt can be kept down so long as the output is low. Transport costs between Anywhere and Nowhere also limit specialisation by adding to the

cost, and therefore taking from the profit, of exchanging salt and sugar. It may happen also that Anywhere cannot spare enough sugar, after meeting all its own requirements, to meet the needs of Nowhere. Or some of the sugar produced in Nowhere may be of a special quality which enables it to find a market in the teeth of competition from Anywhere. In the real world, at all events, specialisation is far from complete. In addition to its export industries, each country has industries which compete with imports, and industries which, like building, transport and distribution, supply a sheltered local market, neither exporting nor competing with imports.

**Extension of the Argument to more Commodities than Two.**

The general lines of reasoning given above can be extended to any number of commodities and to any number of countries. Each country can arrange its products in descending order of comparative advantage—those which it is particularly well suited to produce coming at the head of the list and being exported; those which it cannot so easily produce coming at the bottom and being imported. Where the line of division between exports and imports is drawn depends upon the country's terms of trade—that is, upon the price which it obtains for its exports in relation to the price which it pays for its imports. The more favourable its terms of trade, the fewer goods will the country require to export in payment for a given quantity of imports; while a deterioration in its terms of trade will force it to export a larger quantity and a wider range of goods. The line of division between exports and imports, therefore, is not fixed, but moves up or down as the terms of trade move in favour of, or against, the country.

**The Significance of the Terms of Trade.**

A change in the terms on which a country trades with the rest of the world has repercussions of enormous importance upon its whole economic life. If, for example, its exports fall in price, without any accompanying fall in their real cost of production, the gain which the country derives from trade, and with it the general level of incomes, is automatically reduced. If 100 units of sugar, instead of buying 200 units of salt, buy only 100 units, then wages in Anywhere will fall by 50 per cent. In the same way, if Australian wool and wheat fetch lower prices on the world market, Australian farmers will earn lower incomes and pay lower wages; and since incomes and wages in one industry cannot get far out of line with incomes and wages in other industries, the fall in agriculture will initiate a general reduction. If there are other industries in which Australia enjoys a comparative advantage not much inferior to her comparative advantage in wool and wheat, then these industries will come forward as wages fall and will supplement Australian exports. If there are no such industries, if Australia has a very marked comparative advantage in wool and wheat only, and if both of these commodities are in inelastic

demand by the rest of the world,[1] then Australia will be particularly vulnerable to an adverse movement in its terms of trade. It will gain enormously from trade so long as wool and wheat prices are high, and lose proportionately when wool and wheat prices fall. A particularly striking example is that of Chile, which was until recently extremely dependent upon its exports of nitre and which suffered heavily when artificial nitrates came on the market. The "resourcefulness of supply"—to use a phrase of Marshall's—was low and the vulnerability of Chilean trade was correspondingly high.

Changes in Britain's terms of trade amount, broadly speaking, to changes in the price of manufactured goods relatively to the price of foodstuffs and raw materials. Such changes have been on a large scale, both in the long period and in the short. Within the limits of the trade cycle, for example, the terms of trade tend to move in our favour during the slump and against us during the boom, since the price of foodstuffs and of raw materials oscillates more violently than the price of manufactured goods. Thus we are protected from some of the worst consequences of the slump by a fall in the price of our imports, while our prosperity in the boom is moderated by a rise in their price. In the long period we have had the benefit of steadily improving terms of trade—an improvement which has done more than anything else to raise the standard of living in Britain. It is not because we work harder that we enjoy comforts which our grandfathers could not afford; it is not primarily because our technique is much superior to theirs; it is chiefly because the revolution in transport and the opening up of distant countries (which Victorian capital financed) are bearing fruit in cheap foodstuffs and raw materials. What we can now import with the labour of an 8-hour day would probably have cost us sixty years ago a 10-hour day and a half.

One consequence of the change in our terms of trade has been the eclipse of British agriculture. We import food because we are at a comparative disadvantage in producing it; the cheaper food becomes, the greater is this comparative disadvantage. By employing workers to make manufactured exports we can obtain in return far more food than the same men could produce on British farms. We can afford also to pay them higher wages. By taking men from the land we take them from an occupation where their productivity is low, and by putting them in industry we put them in occupations where their productivity is comparatively high. It was by getting workers off the land that the Victorians were able to raise the standard of living. Had they failed to develop new outlets for labour in industry, the law of diminishing returns would soon have made havoc of Victorian "progress."

**Extension of the Argument to more Countries than Two.**

The fact that there are several countries, not just two, introduces

[1] Notice the qualification; "by the rest of the world." The demand for wheat is highly inelastic but the demand for *Australian* wheat is not.

no important complications. From the point of view of each country the rest of the world is a unit which, for trading purposes, can legitimately be regarded as a single country. It is only when governments begin to manœuvre for bargaining power, or to force trade into the channels of a preconceived policy, that we need to take account of third parties. A policy of Imperial Preference, for example, involves discrimination between empire and foreign countries with a view to fostering trade within the empire. In the same way, a trade agreement may aim at extending trade between the signatories at the expense of other countries. Trade becomes canalised between two countries instead of being carried on, on equal terms, between every trading country. The markets of one country are tied to the export industries of another, while third parties are excluded from markets which they formerly supplied. International trade, as private bargaining between traders gives way to direct bargaining between governments, has to be thought of more and more in terms of international diplomacy—not as a series of bargains struck by free and impersonal competition between traders, but as a struggle for economic power between rival governments. We have not, however, gone so far along this road that the principles of comparative advantage are no longer applicable. Nor, since the principles of government bargaining are a perpetual mystery, are there any other principles to put in their place.

**Trade is Multilateral.**

Although the division of the rest of the world into separate countries works no great change in the principles of international trade, it does create much misunderstanding. People speak as if it were necessary or fair that each country should buy from us as much as we buy from it. But clearly, all that is necessary is that our *total* exports should be as great as our *total* imports—apart from international borrowing or lending. We can have an unfavourable balance with one country so long as we have a favourable balance with some other country. We can go on buying eggs from Denmark, for example, so long as we sell machinery to South Africa. For the South Africans can pay for our eggs by sending oranges to Denmark; or payment may be even more roundabout, South African oranges being sold to France, and French wines going to Denmark in payment. Trade is not bilateral; it is multilateral. A coalminer who buys bread does not insist that the baker must buy his coal, nor does the baker try to pay his rent in bread. In the same way, there is no reason why a country should bully each of its neighbours into buying from it as much as they sell to it; it would generally do much better to give thanks for receiving so much at so little cost.[1]

[1] *cf.* Hume; "And no doubt had the Heptarchy persisted in England, the legislature of each state had been continually alarmed by the fear of a wrong balance; and as it is probable that the mutual hatred of these states would have been extremely violent on account of their close neighbourhood, they would have loaded and oppressed all commerce by a jealous and superfluous caution." (Essay on the Balance of Trade.)

**An Alternative Approach.**

Our exposition so far has been based on the traditional assumption that there is only one factor of production in each country. This assumption is convenient and illuminating, since it simplifies the reasoning and leads to conclusions of great practical value. But it is also unrealistic and leaves us rather in the dark as to the *origin* of a comparative advantage. It is worth while, therefore, to turn back a little and follow a different track, attaching due importance to the existence of a multiplicity of factors and sub-factors of production.

**International Trade and Diversity of Resources.**

International trade, like other kinds of trade, originates in the scarcity and diversity of productive resources. The resources of each country are limited, and they differ from the resources of other countries. By trading, therefore, a country enlarges the range of its resources, drawing on the bounty of other countries instead of contenting itself with home-produced articles. By refusing to trade and seeking after self-sufficiency it denies itself access to resources which are relatively scarce—or cannot be found at all—within its borders; it deliberately restricts its "Lebensraum" or "living-room." That this is so of raw materials is generally agreed; everyone sees the advantage of importing essential raw materials. But what is true of natural resources is equally true of labour and of capital. If one country is well endowed with skilled labour, while in another skilled labour is relatively scarce and unskilled labour plentiful, then it will be to the mutual advantage of both countries if the first sells commodities embodying much skilled labour in return for commodities embodying much unskilled labour. Similarly, countries in which capital is scarce relatively to labour can trade profitably with countries in which labour is scarce relatively to capital. It is differences in *relative* scarcity which gives rise to trade, and these differences are by no means confined to raw materials.

Consider, for example, the international trade of Great Britain. Britain is a thickly populated country with abundant skilled labour and capital and comparatively little fertile land. Britain, therefore, finds it worth while—enormously worth while—to specialise in the production of manufactured articles which make full use of its skill and capital, and to leave food production to countries like Canada, Australia and Argentine in which land is abundant and industrial skill and experience relatively scarce. By exporting machinery, high-quality manufactured articles, and the services of its banking houses, insurance companies and mercantile marine, it receives in return foodstuffs which it could not produce at ten times the cost on its own soil. Not that agriculture in Britain or industry in Australia is wiped out. Distance affords adequate protection to perishable and bulky products, such as milk and vegetables, and gives partial protection to others, such as fresh meat. In the same way, many of the lighter industries in Australia are sufficiently insulated by distance from foreign competition to thrive naturally,

without protection from the State. But it is not only distance which enables British agriculture to survive. There is also an adaptation of farming methods to the comparative scarcity of land. Since land is scarce relatively to labour and capital, it is also relatively dear. There is every incentive, therefore, to economise land and use a high proportion of labour and capital per acre; in other words, to farm intensively. In Canada and Australia, on the other hand, land is relatively abundant and therefore relatively cheap. As a result, farming is extensive, with a low output per acre and a high output per man.

It must not be supposed, however, that Britain's trade is simply a bartering of manufactured goods for foodstuffs. We have always had an extensive trade with other manufacturing countries such as Germany and the United States. For many years we have imported more from the United States than from any other country, and have found in the United States one of our most important markets. In 1913 we imported more from Germany than from any other country except India, and sold to Germany twice as much as we sold to Australia. Part of this trade with other manufacturing countries is in raw materials—cotton imported from the United States, coal exported to Germany. But a large proportion is in finished goods in the manufacture of which one country or another is particularly proficient. No manufacturing country has workers of optimum skill in all industries; some countries excel in the manufacture of chemicals, some in textiles, some in glassware, some in engineering and shipbuilding. Thus within the range of skills at each country's command there are, so to speak, interstices into which trade fits the special skills of other countries. It is not only in the relative scarcity of land, labour and capital that countries differ, but in the relative scarcity of types of land, types of labour, and types of capital equipment. These differences, equally with the broader differences, give rise to trade of mutual advantage to both parties.

To summarise the argument so far. Wherever some factor or sub-factor of production is relatively abundant, it is also relatively cheap. Goods into the manufacture of which that factor or sub-factor enters largely will also be relatively cheap, especially as methods of production will be adapted so as to make extensive use of the cheap factor. On the other hand, goods into the manufacture of which other and relatively scarce factors enter largely will be relatively dear. There will be a tendency, therefore, for the first set of goods to be exported and for the second set to be imported. The resulting exchange of goods will be of advantage because the imports relieve the pressure on the supply of relatively scarce factors, while the exports will provide a vent for the relatively abundant factor.

This argument can now be amplified. First, it should be observed that it is based upon differences in the comparative cost of the factors of production in the trading countries. If the absolute cost of each factor is the same in all countries, the incentive to trade disappears.

For the object of trade is to take advantage of the *comparative* cheapness of commodities in other countries; and if the factors of production are everywhere paid alike, costs of production will be the same everywhere and no commodity will be cheaper in one country than in another. If, for example, there is only one factor of production, labour, which is equally efficient and equally versatile in every country, then no worker will have any reason for trading with other workers either in his own country or abroad. It is only when workers differ from one another, or (what comes to much the same) when there are several factors of production available in different proportions in different places, that trade is worth while. For then there are differences in the relative costs of the factors of production, and differences in the relative prices of their products. Some goods are relatively cheap in one country and some relatively cheap in others, so that there is an all-round advantage in exchanging one set of goods for the other. The volume of trade increases until—setting aside transport costs—prices are brought into line in all countries.

But the earnings of the factors of production are not brought into line. Trade does not make wages equal in all countries, nor interest rates, nor rents. In India, for example, labour continues to earn low wages (even in relation to its low efficiency) because of its abundance relatively to the capital and natural resources of the country. In Britain, capital continues to earn a comparatively low return because there is much more of it per worker than in most other countries. In the leading manufacturing countries, labour earns comparatively high wages because it is skilled and plentifully supplied with capital in a world in which skill and capital are relatively scarce. If the factors of production moved freely from one country to another, these differences could not persist. Workers would emigrate from countries in which wages were low to countries in which wages were high, and capital would flow out to countries in which it is relatively scarce, until the return to labour and capital was equal all over the globe. Since international mobility is low, these movements do not in fact take place. Each country continues to have a relative abundance of some factors and a relative scarcity of others, so that the first factors earn less than elsewhere and the second factors more.

Although trade does not equalise the earnings of the factors of production in different countries, it does tend to level out differences. In the absence of trade, there would be a greater superfluity of labour and a greater scarcity of capital than ever; wages would fall still lower and interest rates rise still higher. Trade has the effect of bringing wages nearer to the level in other countries. The pace is set by the export industries which, having the advantage of cheap, unskilled labour, are able to undersell competitors abroad and, by drawing more and more workers from domestic industries, lever up wages all round. This process can be seen at work quite clearly in Japan where the demand for exports has created a corresponding demand for labour in the export industries and has brought wages

nearer to the level in foreign countries. The movement of goods between countries is, in fact, a substitute for the movement of factors of production. Instead of making use of Japanese labour by allowing it to immigrate, people make use of it at long range by buying its products. Thus they admit it to competition with their own labour only under the handicap of distance. In the same way, unable to borrow British capital for their own industries, other countries can make use of it indirectly by buying British exports in the manufacture of which a great deal of capital has been used. Wages in Japan are raised—although not as much as they would be by direct emigration; and the return on British capital is maintained—although not perhaps so successfully as it would be by investment abroad.

The illogicality of refusing to trade with a country because its wage-rates are low should now be plain. By such a refusal we merely force the country to pay still lower wages; we *intensify* the comparative superfluity of labour in the country by cutting off one source of demand. If there were too many workers in the mining industry and their wages were correspondingly low, we should be doing their cause little good if we stopped buying coal in protest against their low wages; exactly the same is true of workers in foreign industries. Nor are we doing ourselves a service. We do not "protect our standard of living"; we lower it. If a foreign country finds itself able to supply us with iron ore at less than the cost of mining it in Britain, there is no reason in the world why we should refuse to import it unless our ironminers can find no alternative employment or unless ironmining is an industry which we are determined, in the national interests, to preserve. In exactly the same way, if cheap textile products are offered to us—whether they are made by high-paid or low-paid workers makes not one jot of difference—there is no reason in the world why we should exclude them from our markets unless our textile workers cannot hope to be re-employed in other industries. The question is not really one of low wages in foreign countries at all; it is one of potential unemployment in Britain. Whatever the reason why cheap imports are offered to us, we cannot exclude them just because they are cheap (or we should never import anything); nor can we exclude them because some particular factor of production—whether labour, land or capital—is underpaid by our standards; our objection must be rooted in the economic position *within* Great Britain,[1] in some disadvantage to us (such as unemployment) which outweighs the immediate advantage of cheapness.

It is worth adding, finally, that it is not always certain that Britain would gain through a rise in wages in countries like Japan. If Japanese workers spend little on imports, while Japanese manufacturers use a large proportion of their profits for the purchase of machinery and

[1] Setting aside, in this context, political or moral scruples; for example, economic sanctions might be applied against aggressor countries in the form of an embargo on their exports.

other capital goods from abroad, then a change in the distribution of income in favour of labour may have even larger repercussions on the market for imports in Japan than on Japan's power to compete in foreign markets. If the Japanese market is more valuable to Great Britain than the markets in which she has to face competition from Japanese exporters, then she may lose rather than gain from a rise in wages in Japan.

It must not be inferred from this argument that every increase in competition from abroad is to our interest. Whether we gain or lose depends upon the kind of resources brought into competition with ours. Some of our resources—skilled labour, for example—are relatively scarce in other parts of the world and relatively plentiful in Britain. Thus their earnings are high, and will remain high so long as foreigners wish to make use of their services. But if these scarce factors become more plentiful elsewhere—if foreigners acquire skill, accumulate capital, or discover new deposits of coal—our comparative advantage in having large supplies—of skill, capital or coal—is immediately reduced. On the other hand, if the resources in which Britain is most deficient become more plentiful elsewhere—if more territory is opened up, or better methods of farming are introduced—Britain will be able to trade more advantageously. An increase in foreign competition, therefore, is to our interest when it is in commodities which, because of our comparative deficiency in the necessary resources, we are relatively ill-adapted to produce. An increase in foreign competition is contrary to our interest when it is in commodities which, because of our comparative wealth in the necessary resources, we are particularly well-suited to produce. But since we are generally powerless to prevent such increased competition, we must just make the best of it. The competition is felt, as a rule, not in the British market but in our export markets where we have no control; and if it does begin to reach the home market, we should only make matters worse by shutting it out.

It is clear that different sections of the community are differently affected by foreign competition. Those factors of production which are relatively plentiful gain most, since it is their products to which a larger market is opened; those factors which are relatively scarce gain least, or lose absolutely, by being brought into competition with the comparative surplus available in other countries. British capitalists, for example, gain enormously from foreign trade; British landowners would gain even more enormously if foreign trade were suddenly to become impossible. The advantage of foreign trade to the community as a whole, however, is unmistakable. It can be measured in terms of the loss which the country would suffer if it had to exist exclusively on home-produced goods—that is, by the loss inflicted through the diversion of productive resources from the manufacture of exports to the manufacture of substitutes for imports. For most countries, this loss would be very considerable, and the gain which they derive from trade is correspondingly great—so great, one may safely conclude, that it will be rare for some section of the

community not to participate in it. Landlords, who have been cited as losing from foreign trade, are in a highly exceptional position. For agricultural land is not only indispensable to the economic life of every community, but is strictly limited in supply; so that any thickly-populated country which tries to be self-supporting is at the mercy of its landlords, and is forced to pay higher rents than under free competition with foreign agriculture. Unskilled labour is a more typical example. Unskilled workers in Britain share in the advantages of international trade equally with skilled workers. For although in comparison with other countries they are relatively scarce—and might seem to be threatened, therefore, by foreign competition—they would cease to be so scarce if trade with other countries stopped and the demand for skilled export workers vanished. The export trades widen the market for skill and draw off workers who would otherwise compete with unskilled labour. At the same time, the cheap imported foodstuffs which our exports procure lower the cost of living for everyone, and particularly for those who, like unskilled workers, spend a large proportion of their income on food. Not all countries are so dependent on foreign trade as Great Britain. But the rule would appear to be generally applicable—that while some sectional gains from foreign trade are greater than others, sectional losses are rare.

# PART VI—MONEY AND BANKING

## CHAPTER 23

## MONEY

### Money as a Medium of Exchange.

MEN will specialise, as we saw in Chapter 4, only if the limited range of goods which they produce can be exchanged quickly and conveniently for the much wider variety of goods which they wish to consume. But exchange is quick and convenient only if some commodity is singled out as the standard medium of exchange—that is, as money. The alternative of direct barter leads to endless delays and inconvenience. First, there has to be what is called "a double coincidence of wants." The man who wishes to exchange his pig for a gramophone must not only find a buyer for his pig, he must find a buyer with a gramophone to sell. Secondly, the gramophone must be such a good one that he is willing to give his pig for it, and the owner of the gramophone must think highly enough of the pig to be willing to take it in exchange. It is impossible to sell part of the pig or part of the gramophone; both must be offered as indivisible units. Similar difficulties arise when services have to be bartered; the carpenter, the lecturer and the dentist, for example, are likely to have the greatest difficulty in striking satisfactory bargains. The use of money overcomes these difficulties. It allows specialists to obtain for their products or services general purchasing power. Any kind of product or service can be sold for money, and the money so obtained can be used in payment for other products and services. There is an automatic coincidence of wants since everyone wants money; and there is automatic divisibility of the means of payment since monetary units can be made as small as we wish.

The use of money, therefore, arises out of the inconveniences of direct barter in a community of specialists. A good monetary system, by facilitating exchange between producers allows them to carry specialisation further, just as a good transport system, by facilitating exchange between different regions and countries, extends the area of dealings and promotes regional specialisation.

**Money as a Store of Value.**

It is a short step from using money as a means of payment to holding money in reserve against *future* payments. Some of these payments may have to be made in the very near future—when we go shopping, for example. We do not wish to be without money when we board an omnibus, or when we see a bargain in a shop-

window, or when someone calls to collect the rent. Other payments may be more distant. There will be various bills to be met; or we may be saving up in order to buy a motor-car or a house. In addition to the payments which we already contemplate there will be some against which we must guard but which cannot be foreseen—payments due to unexpected illness, or to sudden requests for money from relatives, or to miscalculation of the patience of our creditors. We will generally keep a margin of cash in excess of prospective payments in order to provide for these unforeseen contingencies. Again, since it is in money that our incomes are paid, we may not trouble, because of the expense or inconvenience involved, to invest what we feel to be surplus to our requirements. Or we may decide to hold a store of money in preference to alternative investments, such as bonds, shares or property, because we do not feel confident in the stability of the value of these investments or because we take the view that their value will fall. From all these motives we try to keep by us a stock of money as a reserve of liquid purchasing power. We want money to hold, not just for paying away immediately

**A Definition of Money.**

What distinguishes money from other commodities, either as a medium of exchange or as a store of value, is its acceptability, or—to use a term introduced in the last chapter—its liquidity. Money is, in fact, anything which, by custom or law, is generally acceptable without question in payment for goods and services or in final settlement of a debt. We accept other commodities for their own sake because we wish to consume them; if we do not wish to consume them, we have to go to the trouble of finding a buyer for them. Money, on the other hand, we accept not necessarily for its own sake, but because we know that other people will accept it; we know that there is never any trouble in finding a "buyer" for money—people will take it in exchange for goods of any kind at almost any time. We can never be stranded with money as we can be stranded with other commodities. The man who has just lost his return ticket, for example, will have much less difficulty in inducing the ticket collector to accept a pound note than to accept a book which originally cost £1. The pound note gives him a claim to any goods that he may care to buy; it gives him freedom of choice and security in emergencies. But the book is of use to him only amongst people who know its value and wish to possess it. The pound note is liquid; the book is not.

**Money as a Measure of Value.**

Money is not only the *thing* with which we pay for goods and services; it is also the *scale* in which we measure their value. Money provides a scale of pounds, shillings and pence by which we can make comparisons of value, just as a thermometer provides a scale of degrees centigrade by which we can make comparisons of

temperature. Even if we had to resort to barter some such scale would be indispensable. We could not possibly carry in our minds a complete catalogue of exchange-values, expressing the value of each commodity in terms of a different amount of every other commodity. We should want some common denominator of exchange-values and this common denominator would be money. It would not be money proper—the thing which we exchange. It would be money-of-account—the abstract unit in which price-quotations are made and in which debts are expressed. In Great Britain the pound note is money proper, while the pound sterling is the money-of-account. No one has ever seen a pound sterling—any more than anyone has ever seen a degree centigrade or a mile or any other unit of measurement. What we do see, from time to time, are Bank of England pound notes, with which undertakings to pay one pound sterling may be discharged. The pound sterling does not exist; it is an instrument of thought to assist us in making comparisons of value. The pound note, on the other hand, does exist; it is a piece of paper, fairly constant in value, and generally acceptable in payment for goods priced at one pound sterling. The pound sterling is very ancient; it goes back to Charlemagne. The Bank of England one pound note, on the other hand, dates only from 1928. Thus a new kind of money-proper may come into use without any change in the money-of-account. In the German inflation of 1922–23, the reverse took place. Because of the rapid and incalculable depreciation of the mark, which sometimes lost half of its value over the week-end, contracts were entered into in terms of the United States dollar, the Swiss franc, or commodities like coal and oil, but payment was ultimately made in marks at whatever value they had sunk to. The mark continued to fulfil its function as money-proper, but ceased—in wholesale transactions at any rate—to be the money-of-account.

As a measure of value money makes possible a system of prices. Exchange-values, instead of varying between one transaction and another, become co-ordinated and regularised into market prices. A " market " for each commodity comes into existence and the gain from sales and purchases becomes more calculable. The existence of the market, and the greater calculability which it permits, encourage increased trading, and with this increased trading goes increased specialisation. Once again the use of money lies at the root of specialised production.

**Money as a Standard of Deferred Payments.**

Besides being the unit in which we measure prices (present payments), money is also the unit in which we measure debts (future or *deferred* payments). A debt is an undertaking to pay something—usually money—at some future time; and what has to be paid is measured in terms of some standard unit—the money-of-account. In Britain, for example, contracts of debt are expressed in terms of the pound sterling, but what is actually paid is hard cash—either banknotes or a cheque.

Just as dealings in commodities would be exceedingly complicated without a measure of value, so borrowing and lending would be exceedingly complicated without a unit for the measurement of debts. Instead of borrowing general purchasing power and buying with it such goods as we required, we should be forced to hire or rent the desired selection of goods, returning them (or similar goods) at the end of some agreed period, together with an additional payment of goods as rent. We should be unable to reduce the goods borrowed to a common measure and make repayment in *other* goods of equal value or in a single generally acceptable commodity. It would be impossible also to set debts to a given value against credits to the same value and so to cancel out payments between different traders. Every transaction would stand by itself. Dealings in debts—the work of the Stock Exchange, of the Investment Trusts, of the Banks and other financial intermediaries—would cease.

By providing a unit in which debts can be measured, money makes borrowing and lending enormously easier. A capital market can come into existence in which debts are traded just as goods are traded in commodity markets. With the capital market come large-scale production and increased specialisation. Here, therefore, is yet another link between money and specialisation.

**The Importance of Monetary Stability.**

Money is almost indispensable to economic planning. It is in monetary units that we plan our expenditure, comparing the values of alternative purchases. It is in monetary units that our incomes and expenditure are measured and in monetary units, therefore, that we strike a balance between them, planning how much to spend and how much to save. It is in monetary units also that producers compare the relative costliness of different factors of production, planning methods of production which will substitute relatively cheap for relatively dear factors. Finally it is in monetary units that producers plan their future volume of output, forming their expectations of profit by comparison of present with prospective money-prices. It is of enormous importance, therefore, that monetary units should be as stable as possible. Without reasonably stable units in which to plan, we cannot hope to plan accurately. It would be a curious world if all measurements had to be made in inches which varied in length from day to day. It would be an equally curious economic world if we could not count on fair stability in the value of money from day to day. The only reason why we use a money-of-account is that it provides us with a stable unit of measurement and so allows us to plan ahead. Destroy the stability and people either cease to plan ahead or begin to look for some other money-of-account (as they did in Germany, for example, in 1923). Even comparatively small fluctuations in the value of money may create a state of uncertainty damaging to sound enterprise. Our forecasts of the future may be thrown into confusion because the monetary units in which these forecasts are expressed change unexpectedly in value.

Producers are likely to find, when the value of money falls—and when prices, therefore, go up—that they have made a larger profit than they expected. When the value of money rises—that is, when prices fall—they are likely to make equally unexpected losses. Neither the profits nor the losses can be foreseen so long as changes in the value of money are unforeseen. Fluctuations in the value of money, therefore, create additional uncertainty and make the already difficult task of planning still more difficult. If such fluctuations could be avoided, so that we were able to plan with a stable and dependable measuring rod, economic activity would be steadier and more wholesome, and one of the main causes of speculative booms and chronic depressions would be removed.

**Kinds of Money.**

Although nowadays money is generally accepted for its own sake, the commodities which served as money in early times had nearly all a value of their own, independent of their use as money. Frequently the commodity was a common article of trade like tobacco, tea, salt, skins, olive oil, etc.; amongst pastoral tribes, cattle, sheep and goats were often used. Since these commodities were dealt in extensively in the communities which used them as money they were generally acceptable at, or near, their ordinary market value; they enjoyed a wider circulation, so to speak, than other commodities. But their acceptability was limited. For example, they were not uniform, as pound notes are uniform; it might be necessary where goats were in use as money—as they are to this day in Africa—to call in an arbitrator to decide whether an animal had fallen so far below the customary standard of scragginess that it had ceased to pass as current coin of the realm. Some of the commodities in use as money were not very durable, others not very portable, others again not readily divisible. The only commodities which were free from these disadvantages were the precious metals, which had the additional advantage of being remarkably stable in value. The annual output from the mines was small in proportion to the existing stock and was needed to meet the growing requirements of trade. Over a long period of time, fluctuations in their value did take place, but in shorter periods, between one decade and another, these fluctuations were comparatively trifling. This stability of value added greatly to the popularity of the precious metals as a store of value (since there was less danger of holding a depreciating asset), and even more as a standard of value (since prices were less liable to fluctuate when measured in gold or silver). Once they came to be coined, the convenience of using the precious metals as money was tremendously reinforced. They acquired an almost unchallenged prestige, circulating all over Europe for more than a thousand years with little competition from other kinds of money.

**Coins.**

The earliest English coin—for centuries almost the only coin in use—was the silver penny, a pound weight of silver being divided into

240 pennies. Gold coins, which circulated extensively in Roman times, practically disappeared from use except in large commercial transactions; they were much too dear for the common people to afford in retail trade, but were popular with the merchants, who used them mainly in trade with foreign countries. In the seventeenth century gold coins began to circulate more widely, and in the following century gradually displaced silver as the standard money. At the same time, however, gold coins were themselves being superseded by silver ones.

At first coins were issued by privately owned local mints which made a handsome profit by exacting a stiff charge (known as seignorage) for minting silver bullion. Gradually, however, the private mints were forced to give up their privileges, until nowadays the work of mintage is undertaken exclusively by the State. Owing to the poor technique in use,[1] coins were originally very unequal in weight and shape, and the currency was constantly debased by the issue of counterfeit and clipped coins which circulated side by side with new coins from the Mint. There were always in circulation coins of many denominations, shapes and sizes, some spurious, some of inferior fineness, some sweated, some clipped, some minted in the distant past, large numbers coined in a wide variety of foreign mints. To add to the confusion the King repeatedly depreciated the currency (that is, issued coins to a greater value from a given weight of metal) whenever he was in need of money and had heavy payments to make.

It was not until the eighteenth century that the coinage was in a reasonably satisfactory state. Meanwhile the importance of the coinage was rapidly diminishing as payments came to be made in notes or by cheque.

**Token Money.**

The value of coins was originally equal to their value as metal. What a pound of silver could buy, 240 full-weight silver pennies could also buy. Nowadays, however, we use coins only as tokens. A shilling does not contain a shillingsworth of silver and nickel, nor a penny a pennyworth of copper and tin. Yet both coins are freely accepted at their face value. Why is this? Partly, no doubt, because they are so convenient as a means of payment that we do not stop to ask whether they have any value as metal. But also because they are legal tender. We are legally obliged to accept shillings in payments up to £2 and bronze coins in payments up to 1s.

**Paper Money.**

Coins are now used only for petty cash transactions. In larger payments their place has been taken by banknotes, and in still larger payments by cheques. The use of banknotes in Britain goes back

[1] The coins were cut by hand from a pound weight of silver and flattened and rounded by hammer. Milling was not introduced into England until 1663.

to the seventeenth century, when the landowners began to keep their cash with the goldsmiths, partly out of concern for its safety (as in the Civil War), partly for the sake of the interest offered on deposits. A certificate, or a number of certificates, to the value of the sum deposited was issued by the goldsmiths and these certificates entitled the depositors to withdraw their cash and valuables whenever they wished. Like modern banknotes the certificates were promises to pay on demand; unlike modern banknotes they were receipts made out to a named person and payable only to him or to his order. Gradually, however, banknotes came to be made out to bearer, so that they no longer required endorsement and could pass freely from hand to hand. By the time when, in 1729, printed notes made their appearance it had become the common practice to have notes made out to a named payee or bearer, and by the end of the century notes payable simply to bearer had begun to circulate. But the older practice did not finally disappear until the Bank of England abandoned it in 1855.[1]

The receipts issued by the goldsmiths represented real value, hence banknotes are sometimes spoken of as "representative money." But there was no necessary equivalence in value between the notes in circulation and the gold and other valuables stored with the goldsmith bankers. For the goldsmith, finding that their depositors did not all withdraw their funds at the same time, were able to lend out some of the gold deposited in their custody, or to make loans with notes against which they held no reserve of gold. They could *create* paper money and lend it without preliminary thrift. Modern banks also, like the goldsmiths, issue notes in excess of their gold holdings; part of the note issue, in Britain and elsewhere, is "fiduciary"—based on public confidence, not on a gold backing. But it does not rest with modern banks, as it rested with the goldsmith, to determine how large this fiduciary issue shall be. The law lays down strict limits—sometimes, as in Britain, a fixed *total*; sometimes, as in France and the United States, a fixed *proportion* of the gold held by the note-issuing banks.

It might seem as if the power of the goldsmiths to print unlimited promises to pay must have opened an easy road to fortune. But

[1] Cheques came into use simultaneously with banknotes. Some depositors, instead of taking certificates to the value of their deposit, preferred to keep a running account on which they drew as and when they wished. If they had a debt to settle, they gave their creditor a written note instructing their banker to make payment. The creditor might either call for the money or have it put to his account, often with another banker. The use of written notes, which later evolved into cheques, was a natural development from the older practice of the big landlords in employing a steward to take charge of their money and keep their accounts. Just as it had been the duty of the steward, when authorised by his master, to pay the household bills, meet any gaming debts, and make delivery of the purchase price of any land or property that was bought, so now it was the duty of the banker. For a fuller account of the origin both of notes and of cheques see A. E. Feavearyear: "The Pound Sterling," Chapter V.

the possibility that they might suddenly be asked to redeem their promises limited the size of the issue which they could prudently make. Like a spendthrift, they were at the mercy of those who held their I.O.U's. The spendthrift, asked to honour his notes of hand, must sell off his property. The banker, asked to honour *his* notes, has to call in his loans. If he has lent rashly or for long periods, he stands in grave danger of bankruptcy. In issuing notes he turns to account his reputation or credit as a man of wealth, prudence and honour. His promises to pay enjoy more confidence and are more freely accepted than other men's only so long as his credit is above suspicion. If he loses that credit, his notes will be presented for payment in gold and unless he holds a substantial reserve or can obtain gold from other bankers by offering good security, he will have to suspend payments and close his doors.

This assumes, however, that notes are freely convertible into gold, and that notes are not legal tender. These assumptions could justifiably be made of all banknotes, except those issued by the Bank of England, at almost any time up to 1914. But neither can be made at present. One result of leaving the Gold Standard in 1931 was to make Bank of England notes—by that time the only notes issued in England - inconvertible into gold. The Bank of England still promises to pay on demand "one pound sterling" in exchange for its notes, but no matter how many notes are presented for payment, all that their owners can ever hope to obtain from the Bank is an equal number of similar notes, rather crisper, bearing the same promise. We can buy gold in the bullion market if we are very anxious to have some, but we will certainly get none from the Bank of England. Does this mean, then, that the Bank of England is bankrupt? Far from it. Its notes are legal tender. In using these notes to redeem its promises to pay, therefore, the Bank's behaviour is strictly correct. We want the best money the bank can offer and the Bank has given it to us. No money can be more acceptable than legal tender; it is not only acceptable but *must* be accepted when offered in payment for goods or in settlement of a debt. Once the State has laid down that one kind of money is to be legal tender, its acceptance ceases to be merely customary and becomes enforceable at law.

Since Bank of England notes are inconvertible legal tender, it might seem that the Bank is free to issue as many notes as it chooses. It does not require to redeem its notes in gold; indeed, it does not require to redeem them at all. It is free, therefore, from the obstacles which limited the power of eighteenth-century banks to expand their note issue. But there are other obstacles which limit the note-issuing powers of the Bank of England. What these obstacles are we shall see later. The note-issuing powers of the Scottish and Irish banks, whose notes are not legal tender, are still more limited.

The advantages of paper money are so great that it continues to be used even when it is neither convertible into metallic money nor

legal tender, backed by the authority of the State. During the Napoleonic Wars, for example, the only legal tender money in the country consisted of bad silver and copper coin, and people continued to use Bank of England notes even although they were inconvertible; in the provinces, where Bank of England notes did not circulate, people used local banknotes, which were convertible only into Bank of England notes. Paper money grew so popular that when an issue of sovereigns was made in 1816—the first since 1601[1]—the public showed little interest. Bankers who went to the expense of conveying coin into the provinces had to take it back again to London because of the insistence of their customers on notes. The later popularity of the sovereign in England was due to the prohibition by an Act of 1826 of the issue of notes worth less than £5; no £1 notes were issued again until 1914. In Scotland, where £1 notes have always been legal, the greater convenience and economy of paper money was clearly realised and sovereigns were used chiefly as heirlooms.

**Bank Money.**

Many people think of money as currency (i.e., notes and coin). But there is another kind of money of far greater importance as a means of payment—bank-money. Unlike currency, bank-money, or money deposited with a bank, is not a commodity. There are, it is true, ultra-cautious depositors who make periodic visits to their bank to see for themselves that their money is still there, and who are handed banknotes (a commodity) by the teller. But the banknotes which they inspect are not the same thing as their bank account. Banknotes are oblong pieces of paper decorated with a promise to pay whereas a bank account is an entry in a ledger. This entry gives us the right to ask the bank for a specified amount of currency—a right which, if we make all our payments by cheque, we shall never use. The entry also gives us the right—if our balance is held on "current account"—to draw cheques directing the bank to transfer money from our account to some other person's account. So long as the bank remains solvent and our cheques continue to be accepted freely, we can use our bank balances almost as effectively as notes to make a purchase or to settle a debt. Notes will be rather more effective, since there may be a commission to pay on our cheques and since our signature may be unfamiliar. In times of crisis, too, the risk of banking failures may cause people to insist on payment in notes; notes, being legal tender, are absolutely liquid, but a bank account may be frozen if the bank is unable to meet its obligations. In normal times the difference in acceptability or liquidity is small; we may even prefer payment by cheque to payment in banknotes. Chequable bank deposits, therefore (or current accounts), can normally be classed

[1] The gold coins in use in the 17th and 18th centuries were guineas, not sovereigns.

as money.[1] So also may other bank deposits—deposit and savings accounts. Unlike current accounts, these cannot be drawn on at any moment; the depositor requires to give notice of his intention to draw on them and they are less liquid, therefore, than current accounts. But they are so much more liquid than other assets, and are so commonly classed with current accounts when we are mentally contrasting money with alternative investments, that any sharp line of division between current and other accounts would be false to the facts.

If we include savings accounts in money where are we to stop? Is money lent to a savings bank, money? Or money invested in bills of exchange? Or money used for the purchase of Consols? For an answer to these questions we must go back to our definition. Money, we said, was anything generally acceptable without question in payment for goods and services or in final settlement of a debt. Now acceptability—or liquidity—is a matter of degree. A threepenny bit is more acceptable in Scotland than in England, and more acceptable there than the postage stamps which are sometimes used as money. Unless something is much more acceptable than all possible substitutes, we cannot draw a hard and fast line between what is money and what is not. All we can do is to arrange things in order of liquidity and stop at the first big gap that we come to. We might think that bills of exchange and Consols could be ruled out because they are debts, not money. But some kinds of money are simply debts. Pound notes, for example, are promises to pay—acknowledgements of debt; current accounts are debts due to us by our bank. Pound notes, it might also be suggested, are a perpetual debt; the Bank of England is never likely to be asked to redeem its promises. But the same is true of Consols; they, too, are a perpetual debt unlikely ever to be redeemed. What then is the difference? Is it that Consols bear interest while money does not? Not at all; interest-bearing notes were used as money for over a century. The only ultimate difference is one of degree. Consols are a less liquid asset than money; they are not so generally acceptable as a means of payment. Similarly, bills of exchange and savings bank deposits should be excluded; they stand on the other side of the gap that separates money from less liquid assets.

The convenience of using bank-money is so great we could almost manage without currency altogether, receiving our incomes by cheque and paying them away by cheque. No money would ever change hands; we should simply instruct our bankers to reduce our balance

[1] Cheques themselves are not money: they are devices for transferring the ownership of bank-money from one person to another. Suppose, however, that cheques were used several times over, with a new endorsement on each occasion, before being cashed. They would then be used exactly as banknotes were at one time, and as bills of exchange are at present. If they were freely accepted everywhere, endorsed cheques would be classed as money; but if, as is more probable, they circulated only amongst personal acquaintances, they would be classed as money-substitutes.

and increase our creditor's balance correspondingly. With each payment that we made the bank would owe us less and our creditors more; with each payment made to us, the bank would owe us more and our creditors less. The banks would be in debt to all of us and the only money in existence would consist of this common debt —bank balances—which everyone would accept willingly in payment for goods and in settlement of private debts. Money would consist entirely of claims on the banks, transferable by cheque, and all payments would be made by appropriate entries in bank ledgers.

From this day of universal reckoning in bank ledgers we are still some distance away. We do not all go shopping with cheque-books, and we are not all successful in running up bills (to be settled later by cheque) with tradesmen, landlords and others. We pay for most of our purchases in cash. Hence we still need currency. But currency plays a minor part in modern business. A century ago, it was spoken of as "the small change of credit." To-day, with the rapid growth of banking and the cheque-system, the predominance of bank-money is even more marked. Business and Stock Exchange transactions to an enormously greater value than retail purchases are settled by cheque. If traditional conceptions of money had not so great a hold on us, or if the laws which govern the money-supply were less fantastic, we should pay no more attention to regulation of the note-issue than to regulation of the issue of shillings and threepenny bits. We should try instead to regulate bank-money. But ideas about money are slow to change, a justifiable fear of cranks makes us cautious, and we are content with any system so long as it works. Currency regulation, therefore, continues to take up the time of Parliament and the student of monetary theory; a subject already bewildering enough is made still more complicated. Instead of stating the principles and interpreting the facts about money in terms of a single kind of money—bank-money—we are forced to weave other kinds of money into our explanations and to give them a sham prominence. Currency cannot be ignored. Nevertheless in what follows it will be on bank-money (or credit) rather than on currency that we shall concentrate.

## CHAPTER 24

## BANKING

### What is a Bank?

A BANK is a financial intermediary, a dealer in loans and in debts. It borrows from one set of people and lends to another hiring money and hiring it out again. Some banks draw their capital mainly from their shareholders, others mainly from depositors. Some lend mainly to industry, others mainly to governments, central and local. Some deal in short loans, borrowing and lending for

short periods, others deal in long-term loans, borrowing and lending for comparatively long periods. But however the business of individual banks may differ, their essential function is to gather savings together and lend out what they collect.

To this function others may from time to time be added. In Britain, for example, some of the most important services rendered by the joint-stock banks are advisory and book-keeping rather than financial. The banks are willing to act as trustees on our behalf, to collect dividends for us, to advise us on suitable investments for our money—even to make our will and execute it. They undertake to make payments at such times, and to transfer funds to such places as we may require. Thus they make payment easier, since we can send cheques or drafts by post instead of calling in person or sending a messenger; and they also make payment safer, since "crossed" cheques (i.e., cheques on which we have drawn two parallel lines and written "& Co.") can be paid only into somebody's account and are of no use to a thief.

### The Origin of British Banks

**(1) The London Banks.**—The first British bankers were the goldsmiths. Shortly before the Civil War the government made a large issue of full-weight silver coin, the bullion value of which exceeded its face value. By melting down this coin and exporting it, a handsome profit could be made, and the goldsmiths, who had carried on money-changing as a sideline, saw their chance and took it. They organised the collection of coin, offering 4d. per cent. per day for money left with them to be sorted. Naturally the sorted coins which they returned were depreciated and under weight, while the full-weight coins were all melted down for export.[1] From collectors of coin for export the goldsmiths soon developed during the Civil War into collectors of savings for loan. The landlords' money, which they were anxious to have for sorting, now came to them for safety; while government borrowing, which increased from the Civil War onwards, provided a safe outlet for the deposits made with them. By the end of the century the old trade of goldsmith was beginning to be abandoned for the more lucrative one of banking, and the London bankers were performing all, or nearly all, the functions of a modern deposit bank.

**(2) The Bank of England.** The Bank of England owed even more than the private banks to government borrowing. In 1694 the government, which was then engaged in war with France, found itself badly in need of money. After other methods had been tried,

[1] We have here an illustration of the famous law that "bad money drives out good." This law, called Gresham's Law after a famous Elizabethan statesman, is true only so long as the bad money is not so bad that it makes business dealings practically impossible. In the German inflation of 1923, for example, it was the depreciated notes that began to be driven out of circulation by good foreign money.

a plan was adopted for the setting up of a joint-stock bank—the first in Britain—which was to lend £1,200,000 to the government in return for the privilege of incorporation.[1] The bank, which was given the title of the Bank of England, was to carry on an ordinary banking business and was to have the right to issue notes at least up to the value of its capital. The notes issued by the bank were found to be readily acceptable and circulated freely all over the country. They were held in place of a gold reserve by the private banks, which used them in making payments to their depositors, or procured gold with them from the Bank of England whenever necessary. Thus the three features which to this day distinguish the Bank of England were observable at its foundation. It was a bank of issue; not, as it is now, the only note-issuing bank in England, but even then the leading bank of issue. It was the government's bank—not publicly owned or operated, but guided in its policy by other considerations than private profit: keeping the government's cash balance, lending to the government, arranging for the issue of government loans, advising the government on matters of finance. It was also, in embryo, a central bank: leading and directing the monetary policy of the country, and exercising centralised control over the gold reserves which the commercial banks came in time to deposit with it.

(3) **The Country Banks.**—The Bank of England's charter gave it, for over a hundred years, a monopoly of joint-stock banking in England. From 1708 onwards it was illegal for any firm with more than six partners to carry on the business of banking, and only small private banks, therefore, could be founded. It was not possible—as it was in Scotland—for a few strong banks to grow up with branches in all parts of the country. The Bank of England refused to establish country branches and the London bankers could not build up the necessary organisation. Until the Industrial Revolution, therefore, the provinces were left almost entirely without banking facilities.[2]

The first country bankers, who began to make their appearance in the middle of the eighteenth century, were local shopkeepers. At that time remittances were made from one part of the country to another by means of "inland" bills of exchange. These bills were rather like post-dated cheques—that is, they were documents entitling the holder to payment at some future date, the amount to be paid and the time and place of payment being stated on the bill. Bills were

[1] The Bank not only obtained its charter in return for a loan to the Government but had repeatedly to make additional loans as the price of its renewal.

[2] It should perhaps be added that until the Industrial Revolution there was little need for banks in the provinces. The trade of the country centred on London, goods moving in from the provinces and from abroad and moving out to the provinces and to foreign countries. It was from London that the merchants who traded in these goods were financed, and to London that the rents of the big landowners were sent. Until the growth of localised *industry* (as distinct from *commerce*) the needs of the provinces were quite adequately met by the London banks. (Feavearyear: *op. cit.*, page 148).

the means by which London paid for the foodstuffs which it imported from the countryside; the farmer sent his grain and his cattle to the London market and drew a bill for the price on the dealer who bought them. The same bills served as payment for the wines and silks which landowners imported from London; the farmer, looking for someone to buy his bill, found a ready market in the shopkeepers who retailed London goods. As trade with London increased, dealing in bills came to be more profitable than dealing in goods, and the provincial shopkeepers began to blossom into bill-brokers; they bought bills even when they had no remittances to make and undertook to obtain bills even when they had none on hand.

The next step was the development of the inland bill from a convenient method of making remittances into a credit instrument. Manufacturers who wished to borrow money could not go to the banks for an advance as they can nowadays. What they did was to pay for their raw materials with bills drawn for a period of several months. The manufacturer was thus put in possession of working capital by the person who consented to take the bill instead of cash. If the manufacturer's credit was good, he might also sell similar bills directly to the shopkeeper-bankers, and so obtain funds on the strength of a promise to pay, backed by adequate security. Directly or indirectly bills reached the shopkeepers in amounts too great for their personal resources.[1] They began to compete for the savings of the farmers and landlords—whose bills on London they were already buying—and so entered on the business of deposit-banking.

At the same time they began to forward bills for re-sale[2] in the London money market, and so came into touch with the London banks who acted as their agents.[3] It was necessary for them to keep

[1] As inland bills were used as currency (especially in the north, where the country banks did not issue notes) it was not always the banker who provided the manufacturer with credit.

[2] Strictly speaking, for *re-discount*. Since credit is given by the holder of a bill to the person on whom it is drawn, bills fetch less than their face value (i.e., the sum to be paid when the bill becomes due for payment or "matures.") The difference between the price of the bill and its face value represents interest on the loan and is known as discount. To discount a bill is to buy it at less than its face value: to re-discount a bill is to buy it from someone who has already discounted it.

[3] This contact between the country bankers and the London money market was of great importance in transferring capital from the agricultural districts of the South and East to the industrial districts in the North and West during the Industrial Revolution. The landlords and farms of the South and East, finding no employment for their savings, put them on deposit with their banks. The banks used these funds in London to discount bills drawn by the manufacturers of Lancashire and Yorkshire. Much of the working capital of the textile industries, therefore, was provided out of the profits and rents of the landlords and well-to-do farmers of the agricultural districts. Now that the banks have branches in all parts of the country, however, bills are unecessary for the transfer of capital: deposits made in one county can be lent out directly in another.

an account with these London banks, and this account formed an important part of their reserve, since they could draw on it if in need of gold. At the same time the London banks were coming to hold their reserve either in Bank of England notes or in the form of an account with the Bank of England. Thus the reserves of the banking system of the country were concentrated more and more under the control of the Bank, and it was the Bank's gold stock which, as the country's final reserve against notes, had to stand the brunt of a panic.[1]

Such panics were all too frequent. The country banks were constantly over-issuing their notes and lending too freely for their own safety. In a bad year they went down like ninepins, unable because of their smallness and inexperience to weather a major crisis. No fewer than 89 country banks failed in the three years 1814–16, and bankruptcies on an even bigger scale were recorded in other years. But as fast as one bank failed, another sprang up to take its place. At the beginning of the Napoleonic Wars there were 400; by the end there were no less than 900.[2]

The great weakness of the English banking system—in contrast with the Scottish—was the absence of strong joint-stock banks with many branches to spread the risk of failure. But until the Bank of England's monopoly was broken it was impossible to found a joint-stock bank. At last, after a particularly severe crisis in 1825, joint-stock banking *in the provinces* was legalised. An Act of 1826 permitted joint-stock companies to carry on the business of banking and to issue notes, provided that they did not open an office within 65 miles from London. In 1833, after a successful agitation led by Thomas Joplin, the London area, too, was thrown open to joint-stock banking, on condition that banks opening offices in London did not issue notes.[3] The National Provincial Bank was formed in the same year, and the Westminster and the Midland shortly afterwards. Joint-stock banking developed rapidly, partly through natural growth, partly through amalgamation with the private banks. By 1900 only 19 private banks were left, and by 1937 this number had fallen to three.[4] The joint-stock banks, after reaching a maximum of 118 in 1878, began to fall in number as a result of amalgamation between themselves. The amalgamation movement reached its peak at the turn of the century when the London banks, seeking an outlet for their surplus funds, invaded the provinces in search of customers. Between 1895 and 1905 nearly one-half of the total number of banks were absorbed, while in the same period the number of bank branches almost doubled. In 1918–19 there was a further orgy of amalgamation. The public

[1] The above account is summarised from Heavearyear, *op. cit.*, Chapter VII.

[2] In a single year, 1809, a hundred new banks were started.

[3] By this time the only notes in circulation in London were those of the Bank of England, as the private banks there had given up the issue of notes.

[4] Blydenstein & Co., Hoare & Co., and Japhet & Co. The total assets of these banks are under £14,000,000.

grew alarmed at the threat of a banking monopoly, and a Treasury Committee recommended that no further amalgamations should be permitted without the consent of the government. Very few amalgamations, and still fewer of any importance, took place after 1920. The chief—and the only one since 1928—was that of the District Bank with the Country Bank in 1935. This amalgamation brought the number of joint-stock banks in Great Britain down to 26 (including the eight Scottish and three Northern Irish banks, but not including the Bank of England). Of this total, three of the English, four of the Scottish, and two of the Irish banks are no longer independent, and three more can be disregarded since their assets come to less than £5,000,000. This leaves 14 large independent banks, controlling between them over 12,000 branches and sub-branches and assets totalling over £3,000 millions.

**The Scottish Banks.**

In Scotland, where the Bank of England's monopoly did not apply, and the Bank of Scotland's monopoly had been allowed to lapse after 1716, banking developed along healthy lines almost from the start. Strong joint-stock banks were founded with a large capital of their own, and these banks soon began to establish branches all over Scotland. Alongside the joint-stock banks were a number of private banks, with men of real substance—and nearly always more than six of them—as partners. The strength and unity of the Scottish banks, which grew as the private banks were amalgamated, put them in an excellent position to withstand a panic, and, by giving confidence to their depositors and note-holders, made a panic less likely. Bank failures, therefore, were very rare; in the whole history of Scottish banking there has been only one serious failure in which the public was not ultimately paid in full—that of the City of Glasgow Bank in 1878.

The Scottish banks were the first to develop branch banking. The policy was begun by the British Linen Company (founded in 1746), which, in its efforts to foster the linen trade, was naturally led to make advances of credit to linen producers and so drifted into banking. The original object of the company was soon forgotten, but the branches which it had opened in various parts of the country remained. The success of the British Linen Company (now the British Linen Bank) encouraged other banks to follow its example, notably the Commercial Bank, which started 14 branches within 10 years from its foundation in 1810. Fifty years ago there were still nearly half as many bank branches and sub-branches in Scotland as there were in England; and even now, when the ratio is less than 1 : 5, it is much higher than the ratio of population and wealth.

The Scottish banks were also the pioneers of the pound note and the overdraft. Pound notes have been issued by the Bank of Scotland ever since 1704, and have never been illegal in Scotland as they were in England between 1826 and 1914. They provided a reliable and acceptable paper currency at a time when the coinage was a hopeless

jumble, and were of great benefit in economising gold when Scotland was too poor to afford it. The overdraft, at an early date, became almost equally familiar. It originated in the Cash Credit Bond, a contract under which a Scottish bank agrees to lend up to some fixed amount to any borrower who can find two men of substance who will guarantee repayment of the loan. If the borrower is in funds he can repay part of the loan and so save interest, since he is charged interest only on the amount outstanding. In the modern overdraft a similar principle is followed. But instead of requiring the guarantee of two men of substance as security, the bank generally asks for "collateral"—that is, the borrower must pledge Stock Exchange securities, a life insurance policy, the title deeds of his house, or some similar assets to a sufficient value to satisfy the bank. Interest is charged on the outstanding balance and the loan comes up for renewal every six months or a year.

**The Balance Sheet of a Commercial Bank.**

Now that we have seen how banks originated, let us examine more closely what they do. We may begin by running over the main items in a typical bank balance sheet, first from the point of view of the bank liabilities (the money put at its disposal) and then from the point of view of its assets (the use to which that money is put). The combined balance sheet of the 15 English joint-stock banks, given on page 309, will serve to illustrate the main items on both sides and their relative order of magnitude. In other countries, where the banks perform rather different functions, some of these items are called by different names, have a different order of importance, or are replaced by other items which have no counterpart in Britain. These differences will be discussed later when we come to deal with the peculiarities of British banking.

LIABILITIES

A bank's resources are supplied partly by its shareholders and partly by its depositors. The shareholders supply the capital, and this is the first item on the liabilities side; as in all joint-stock companies the capital subscribed by the shareholders is represented as a debt due to them. So also is the next liability, the reserve fund which the bank has built up out of past profits and withheld from its shareholders.[1] The undivided profits, which will later be paid out in dividends, may also be classed as a liability to the shareholders. The next liability—against notes—has disappeared from the balance sheets of the English commercial banks except in respect of notes circulating in the Isle of Man. Notes, being promises to pay legal tender on demand, obviously involve the issuing bank in a liability to the note-holder. Acceptances, which come next, are *contingent* liabilities and are offset by an equal item on the assets side. When

[1] In addition to its published reserves, the bank may hide secret reserves under the heading "other accounts"; or it may enter the value of its property on the assets side of the balance sheet at a specially low figure.

a bank "accepts" a bill of exchange it is guaranteeing a loan; it makes itself responsible for payment of the bill when it falls due, while the drawer of the bill (who obtains credit by selling it) undertakes to put the bank in possession of the amount due to be paid before the bill matures. If all goes well, the bank pays and is paid more or less simultaneously and liability and asset disappear together. The remainder, and the bulk, of the bank's liabilities are to its depositors. First, there are its liabilities on current account; current accounts earn no interest, are withdrawable on demand, and can be drawn on by cheque. Second, there are the bank's liabilities on deposit account; deposit accounts are more of the nature of savings accounts —they earn interest in accordance with a fixed schedule, can be withdrawn only when a specified period of notice has been given to the bank, and are not chequable. The distinction between current and deposit accounts is perhaps better conveyed by the more general terms—"demand deposits" and "time deposits."

*Combined Balance Sheet of the Fifteen English Joint-Stock Banks and of the "Big Five" at December* 31, 1938.[1]

LIABILITIES (£ millions).

| | *15 Banks* | *5 Banks* |
|---|---|---|
| Paid-up Capital | 80·6 | 65·6 |
| Reserve Funds | 61·6 | 50·5 |
| Undivided Profits | 5·8 | 4·4 |
| Notes in Circulation (in Isle of Man) | ·1 | — |
| Acceptances | 132·7 | 104·8 |
| Deposit, Current and Other Accounts | 2,268·8 | 1,950·1 |
| | 2,549·6 | 2,175·4 |

ASSETS (£ millions).

| | *15 Banks* | *5 Banks* |
|---|---|---|
| Cash in hand and at Bank of England | 352·6 | 294·6 |
| Money at call and short notice | 152·1 | 120·7 |
| Bills discounted | 234·4 | 217·2 |
| Advances, Loans, etc. | 977·7 | 855·8 |
| Investments in British Government Securities | 571·7 | 484·2 |
| Other Investments in Bonds, Stocks, etc. | 81·2 | 61·2 |
| Cover for Acceptances | 132·7 | 104·8 |
| Premises and Sundries | 57·3 | 37·1 |
| | 2,549·6 | 2,175·4 |

[1] The 15 banks are :—Barclays, Lloyds, Midland, National Provincial and Westminster (the "Big Five"): Coutts & Co. (owned by the National Provincial), District Bank, Glyn, Mills & Co., Martins Bank, and Williams Deacon's Bank (owned by the Royal Bank of Scotland): Baring Bros & Co., British Mutual Banking Co., Isle of Man Bank, London Merchant Bank, and the Union Bank of Manchester. The first ten banks, together with the National Bank of Ireland, make up the London Clearing Banks; that is, they alone enjoy clearing facilities at the London Clearing House.

## Assets

**(1) Cash in Hand and at Bank of England.**—A bank's assets represent the use which has been found for the money put at its disposal by shareholders and depositors. First, some money has to be kept as cash in hand so as to provide a first-line reserve against withdrawals by depositors. The amount needed depends upon the maximum withdrawal of cash which the bank anticipates at any one time. Every day some money is paid in and some is withdrawn. If the rates of inflow and outflow were equal it would be unnecessary to hold more than a trifling reserve; if they were even regular and predictable, the bank could always provide itself with cash in time and again a small reserve would suffice; but as inflow and outflow are neither regular nor predictable, a fairly large reserve is necessary. In the days when there was a real danger of a run on the bank, a much larger cash balance than at present had to be held, but now that this danger has practically disappeared, some 6 per cent. of deposits is found to be adequate.

Secondly, a balance is kept with the Bank of England as a supplementary reserve. This balance is drawn on whenever one bank finds itself in debt to another. Suppose, for example, that people who bank with the Midland pay into their accounts cheques drawn on Barclays, while depositors with Barclays pay into their account cheques drawn on the Midland. What is really happening is that money is being withdrawn from each bank and paid into the other. If the two sets of cheques are equal in value, the two banks can agree to cancel them against one another at the Clearing House—an institution at which cheques drawn by the customers of one bank on the customers of another are "cleared" or set off against one another.[1] No difficulty arises. But if the first set of cheques (those paid into the Midland) are more valuable than the second set, then it will be necessary for Barclays to pay over the difference. Payment is made by a cheque which is drawn on Barclays account with the Bank of England and paid into the Midland's account. Just as till-money provides for sudden withdrawals of cash, so balances with the Bank of England provide for sudden withdrawals by cheque.

These two items—cash in hand and balances with the Bank of England—form the cash reserves of a joint-stock bank. Together they generally amount to between 10 and 11 per cent. of the bank's deposit liabilities, this proportion being spoken of as the bank's "reserve-ratio." In the combined balance-sheet on page 309 the reserve-ratio appears to be considerably higher than 10 per cent. This is due to the inclusion under "Cash in hand and at Bank of

[1] Informal clearing arrangements existed in the eighteenth century, when the clerks who were sent to collect cheques drawn on other banks began to meet together instead of walking from bank to bank. In 1775 the London Clearing House was set up to clear cheques drawn on the London banks. The Joint-Stock banks were admitted in 1854 and the Bank of England ten years later. At present, eleven banks enjoy clearing facilities, and the value of the cheques cleared is about £5,000 millions per month.

England" of an item generally described as "the float" and consisting of cheques drawn on other banks and in course of collection from them. The "float" is almost as good as cash, except that it includes cheques which may ultimately be refused and returned to drawer. But since each bank is both creditor and debtor, collecting cheques and being collected from, it cannot count on any net balance of cash in its favour once all cheques have been collected.[1]

(2) **Money at Call and Short Notice.**—Call-money is lent on a day-to-day basis, mainly to bill-brokers. The bank manager in charge of "the till" at the Head Office draws up a statement every morning showing the probable movement of funds out of, and into, the bank. If he has reason to expect a credit balance, after allowing a margin for contingencies, he makes loans at call (or for longer periods if he expects the balance over outgoings to continue). If he is faced with a debit balance, he calls in loans. Later in the day, if money comes in unexpectedly, he makes further loans to the money market. He may also, towards the close of business hours, make overnight loans to stockbrokers to enable them to balance their account. All of these loans are callable next day and normally carry a very moderate rate of interest.[2] Other loans, for periods up to fourteen days, appear under the heading "money at short notice." These loans are made either to the money market at seven days' notice, or to the Stock Exchange for the settlement (that is, to allow stock to be taken up in the intervals between the fortnightly settlements).

In normal circumstances, money at call and short notice accounts for about 7 per cent. of the deposits of English clearing banks, half of this consisting of loans at short notice to billbrokers and the other half being equally divided between call money and loans to the Stock Exchange.

(3) **Bills Discounted.**—Rather longer loans, the normal period being up to three months, are made through the purchase of bills of exchange. A bill of exchange is a credit instrument, transferable by endorsement, and requiring one person (the drawee) to pay a certain sum of money to another person (the drawer) at a specified date. On the bill is stated the amount which is to be paid, the time and place of payment, and the names of the drawer and drawee. Some bills are used in commerce to finance purchases of imports from another part of the country (inland bills, now little used in Britain) or from abroad (foreign bills). The importer pays for his goods, not in cash, but with a bill of exchange drawn on his bank or on one of the London acceptance houses (which specialise in this kind of business). The

[1] The "float," together with "premises and sundries," is roughly equal to the total of capital, reserves, and undivided profits. The remaining assets of the bank, therefore, omitting cover for acceptances, are roughly equal to the total deposit liabilities of the bank. If we express each of these assets as a proportion of deposits, as is customary, the proportions sum up to approximately 100 per cent.

[2] In a crisis, however, call money rates are the first to rise.

exporter turns over the bill to his bank (together with various other documents) and asks to have it discounted at once. It is sent for acceptance to London, the acceptance house endorsing it by writing its name on it. This guarantees the bill, and since the credit of the acceptance house is unimpeachable, makes it readily negotiable. After acceptance, the bill is sold (or, to use the technical term, "discounted" and the exporter is then put in possession of his money. The importer, in the meantime, is given credit by the holder of the bill until it falls due, by which time he has presumably disposed of the imported goods and so acquired the funds necessary for paying off the bill. He sends the sum due to the acceptance house which in turns pays the holder of the bill.

Bills of exchange are thus a very liquid investment. They are a self-liquidating advance, since when the goods bought come to be sold, the money originally advanced is automatically recouped; and they are a very short-term advance since they are rarely drawn for periods exceeding three months, and a bank can arrange its portfolio of bills so that some mature every week, or even every day. Hence they are much in demand, the normal holdings of the British banks amounting to about £100 m., but falling considerably below this level when international trade is inactive.

Of even more importance than commercial bills are bills issued by the government. These Treasury Bills, with which the Floating Debt of the government is financed, are tendered for every week by the banks and the discount houses, allotment being made to the highest bidders. In the operations of the money market, Treasury Bills have largely taken the place of bills of exchange, the supply of which has fallen heavily in recent years. The banks have on occasion held over £300 m. in Treasury Bills, this total being reached in 1933 when the supply was swollen in connection with the Conversion Scheme of the previous year.[1] In other years—for example, in 1929—the banks' holdings have fallen as low as £100 m.

There are clearly wide fluctuations in the bill portfolios of the banks. The supply of bills varies with the state of trade and with the fiscal policy of the government. The demand varies with the interest obtainable and with the liquidity of other bank assets. The fluctuations in "Bills Discounted," however, are generally offset in large measure by opposite fluctuations in "cash" and "short loans." In 1934, for example, when the decline in "Bills Discounted" was particularly drastic—from 18·1 per cent. to 12·2 per cent. of total deposits—there was an equally abnormal increase both in "cash" and in "short loans,"—the first increasing from 10·8 per cent. to 11·3 per cent. and the second from 5·2 per cent. to 7·1 per cent.[2] It is, in fact, a point of policy with the Joint-Stock Banks to maintain a fairly steady proportion of about 30:70 between their *total* liquid assets (those so far discussed) and their other assets (those still to be discussed).

[1] R. J. Truptil: "British Banks and the London Money Market," page 98.

[2] These figures are for the London Clearing Banks.

**(4) Advances, Loans, etc.**—Advances are the most important and most profitable of bank assets. Up to 60 per cent., and rarely less than 35 per cent. of a bank's deposits are advanced to its customers, generally on overdraft. Borrowers are allowed to withdraw money up to an agreed limit and are charged interest on the outstanding balance at a rate which is normally 1 per cent. above Bank Rate with a minimum of 5 per cent. ("one over minimum five").[1] The security for the advance varies with the customer. The most usual collateral consists of stock exchange securities, but advances are also made against real estate, commodities, life insurance policies, personal guarantees, or on the unsupported credit of the borrower. Most advances—probably about nine out of ten—are for less than £1,000, many of these small advances being made to professional and private persons, who form more than half of the total number of borrowers. Advances of less than £1,000, although of great importance numerically are of much less importance by value. It is the large advances to industrial concerns—in amounts up to several million pounds—that make up the bulk of bank advances. About half of the total amount lent by the joint-stock banks consists of advances in excess of £10,000.

These advances are renewable, as a rule, every six months. Hence the banks do not look with favour upon the locking up of the money which they lend in any long-term investment. From their point of view, the ideal advance is one which is made to tide over a temporary shortage of funds. At harvest-time, for example, the farmer is faced with a heavy outlay in wages which he will recover when his crop is sold. Instead of holding money throughout the year to meet this seasonal outlay (and so losing the use of a large part of his capital), he can borrow an overdraft from a bank, repaying the loan within a few months. The peak load on his capital, so to speak, is carried by the bank, leaving him with no idle capacity in off-peak months. Since the seasonal peaks in other industries occur at other times of the year, the bank is able to dovetail the borrowings of one industry with the repayments of another and maintain a revolving fund of credit on which each industry can draw in turn. Meanwhile continuous repayment preserves the liquidity of the bank's assets: by reducing new loans and by refusing to renew old ones when they fall due, the bank can cut down advances quickly if it finds itself short of money. Not that it is only seasonal loans which offer these advantages. *Any* loans made for short periods to supplement the working capital of industry and commerce are just as liquid. An industrialist who obtains an advance to cover wage-payments and purchases of raw materials is in a position to make repayments once the finished goods are sold; and a merchant who borrows to finance his holding of stocks realises from sales enough to repay the original advance. The industrialist, if he is producing a steady stream

[1] In exceptional circumstances, some advances are made on special terms at rates below this nominal minimum.

of goods and covering his costs, can constantly renew his indebtedness to the bank as each batch of goods reaches the market. The merchant, if he wishes to maintain his stocks at their previous value, can renew his loan and apply his sales revenue to the replenishment of stocks instead of to the extinction of his debt to the bank. Loans of either kind, like seasonal loans, are self-liquidating, and are in danger of being frozen only if they are renewed when the borrower is failing to cover his costs.

Although the short term, self-liquidating advance is the ideal of the British banker, not all advances answer to this description. British farmers, for example, have borrwed far more from the banks to assist them in buying their farms—amongst the most illiquid of assets—than for the financing of seasonal outlays. Many householders have used a bank advance for the purchase of their house. Even extensions of plant may be carried out with the help of an overdraft. Loans for all of these purposes—not, in the aggregate, a very large proportion of the total—are illiquid unless the borrower has ready access to money at short notice. They are really long-term advances, which all borrowers can be reasonably sure of repaying eventually, but which many would have the greatest difficulty in repaying within six months.

Since the success of a banker depends mainly on his skill in lending, the banks show great solicitude towards their customers. They claim that no reasonable request for an advance is refused, and that, if they are forced to restrict credit, advances are the last of their assets to be reduced. These claims, it is true, seem to be mutually destructive; it is impossible to restrict advances and still grant every reasonable request for one. But a banker's standards of reasonableness are notoriously variable, so that the first claim may well be true, although rather meaningless, like the statement "white mice are white"; while the facts, which point to a real reluctance on the part of the banks to call in their loans, substantiate the second claim. There is no need, moreover, to presuppose any wide departure from normal standards of reasonableness when advances are being reduced. It is natural to be more exigent in a period of credit restriction and falling prices. Again, the number of requests reaching the banks at such times is abnormally low, while during a period of credit expansion a much larger number of applications for loans is submitted. Thus the banks are hardest pressed to lend when most free to do so, and least able to lend when they least wish to. Since the demand for loans fluctuates with the supply, the need to accommodate the one to the other by a policy of rationing is very much reduced. But it should be noticed that it *is* by rationing, rather than by a change in price, that supply and demand are balanced. The rate charged on advances is linked, by custom, with Bank Rate; and although, as it happens, the rate generally increases when demand increases, there is no reason to believe that variations in it are of major importance in limiting or in stimulating the demand. The real balancing factor is the banks' standard of reasonableness, the kind of security which

they demand before granting an advance. This standard is set, partly by the views which the banks hold of trade prospects, but chiefly by the proportion of their deposits which they are willing to advance to customers.

(5) **Investments.**—Of all bank assets, investments in long-dated securities are the least liquid. The securities (mainly British government bonds) are redeemable only after considerable periods of time, some (Consols) being completely irredeemable. As a result, the capital value of the banks' investments is free to fluctuate, rising and falling inversely with the current rate of interest. After 1931, for example, the sharp fall in interest rates was reflected in a rising price for gilt-edged securities, the movement being reversed about the beginning of 1935, when interest rates began to increase.[1] The banks profited from an appreciation in the value of their bonds in the earlier years and made a corresponding capital loss when their holdings subsequently depreciated.

Although bonds are a comparatively illiquid investment, they are a highly marketable one. There are organised dealings on a gigantic scale in British government securities at quoted and steady prices; a large block can be disposed of in a short time at a price which is known in advance. Even discounts are not so marketable; and in fact, since they would bear the bank's name on them once sold, and might give rise to undesirable speculation as to the bank's reasons for selling, discounted bills are never marketed by a British bank, but are invariably held to maturity. Advances cannot be sold for the very good reason that there is no market in them; an advance involves a personal assessment of the borrower's credit—of the security which he offers, of his character and ability, and of the prospects of his business. Gilt-edged securities, on the other hand, require an assessment only of the government's credit, on which opinion is more easily formed, and less likely to differ. There need be less hesitation, too, about reducing investments than about reducing advances. The sale of securities is an impersonal transaction which allows the most willing buyer—a stranger to the bank—to come forward and provide the money which the bank requires. But the calling in of a loan is a very personal transaction. It may mean the loss of a customer; it may even mean the bankruptcy of a customer; and it is quite likely, if too sudden and too pressing, to force the borrower to realise his assets at far below their true value. Thus if a bank finds it necessary to contract credit, it is likely to play for time by selling off investments rather than reduce its advances too rapidly. Similarly, if a bank finds itself accumulating large surplus funds it will generally invest them in government securities.

[1] The fall in interest rates between 1931 and 1935 was itself due in part to heavy purchases of bonds by the Joint Stock Banks. These purchases raised the price of bonds and reduced the yield on them, the rate of interest on new long-term borrowing falling in sympathy.

rather than relax too far the standards of security which it requires from borrowers. The security holdings of the banks, therefore, are largely a residual item, varying inversely with advances. When trade is good, the banks tend to sell off investments to provide for the increasing requirements of their customers. When trade is bad, the banks are repaid some advances and forced to demand repayment of others which seem likely to be frozen. As advances are liquidated, funds become available for investment, and if the risk of depreciation does not seem too great, the banks increase their holdngs of government securities.

These fluctuations between boom and slump are not the only forces which disturb the ratio between bank advances and investments. In war-time, for example, the banks are generally required to support the credit of the government by participating in issues of war loan. This adds to their holdings of securities; and on such a scale that, in spite of the expansion of credit which accompanies a war, investments mount more rapidly than advances. At the end of the war the banks are likely to sell off some of the securities which they have acquired and use the proceeds for making additional advances to industry. Between 1922 and 1930, for example, the proportion of bank deposits invested in long-dated securities fell from one-fourth to one-seventh. After 1930 the trend was reversed, partly by the depression and partly by the fall in long-term rates of interest relatively to the rate on bank advances. The banks maintained their overdraft rate at 5 per cent. even when long-term rates were below 3 per cent. and short-term rates so low that it was actually cheaper to borrow money for three months than to print it. As a result of this policy many businesses paid off their outstanding bank loans out of capital borrowed on long-term on the Stock Exchange, or financed themselves out of their own reserves without recourse to the banks, or borrowed at low rates outside the banking system altogether. The growing importance of holding companies worked in the same direction, for the debts of one company in the group could be financed out of the credit balance of another. The larger and more integrated the business unit, the greater were the chances of such internal cancellation of debit against credit and the less was the need to borrow from outside. Thus there was a steady trend all through the 'thirties towards a lower proportion between advances and bank deposits and a correspondingly higher proportion between investments and deposits. The change was largely concentrated in the years of intense depression between 1930 and 1933, the advances of the London Clearing Banks falling from £994 millions in June, 1930, to £787 millions in June, 1933, and investments increasing from £278 millions to £587 millions over the same period.[1]

[1] In other countries, a similar change took place. The investments of the Canadian Chartered Banks, for example, rose from $526 million in 1929 to $1,314 million in 1936 while advances fell from $1,321 million in 1929 to $725 million in 1936.

Banking was becoming increasingly an investment business, with the government as the chief borrower. In the middle and later 'thirties, in fact, fully half of the deposits of the clearing banks were being lent to the government.

### Peculiarities of British Banking

(1) **Branch Banking.**—In Britain a single bank may have over 2,000 branches. In the United States, by contrast, the *total* number of branches does not exceed 5,000, most of these being sub-offices in the home town.[1] In Britain the system of branch-banking has enabled a small number of very strong banks to build up a semi-monopolistic position, the Big Five controlling five-sixths of total deposits. In the United States, on the other hand, the number of competing banks is enormous (over 15,000), and the 250 largest banks control only one-half of total deposits. Thus the danger of monopoly is greater in Britain and of bank failures in the United States. That the second danger is a very real one is shown plainly in the failure of 10,000 American banks between 1921 and 1933. The first danger is less urgent and would not necessarily be removed if the number of competing banks were larger.

America is the only country of importance in which there is no developed system of branch banking. In Germany there are only three large deposit banks, all of them working under close government control[2]. In France there are four.[3] As in Britain these banks have their head office in the capital city, and branches all over the country. They can promote the flow of funds between places where there is a surplus and places where there is a shortage, are able to put a large centralised reserve behind any point of danger, and by scattering their investments can spread their risks much more successfully than a small local bank. Thus it requires a major crisis —such as brought down the Darmstaedter Bank in 1931—to threaten their stability.

(2) **Short-term Lending.**—The capital of British banks is small relatively to their deposits. For the English joint-stock banks, capital and reserves (including undivided profits) average only 6 per cent. of all liabilities, while their uncalled capital forms an additional reserve of 7 per cent. Working on so slender a foundation of their own resources, and faced with the possibility of sudden withdrawals by their depositors, the banks must be in a position

---

[1] In many of the States branch-banking is prohibited by law and in many others, an *extension* of branch-banking is prohibited. The law is, however, partially circumvented by a system of interlocking directorates, and by chains of independent banks under the control of holding companies.

[2] The Deutsche Bank und Discontogesellschaft: the Dresdner Bank (which owns the Darmstaedter und Nationalbank): and the Commerzbank (which owns the Barmer Bankverein).

[3] The Crédit Lyonnais, the Société Generale the Comptoir Nationale d'Escompte, and the Crédit Industriel et Commerciel.

to turn their assets readily into cash. Hence they confine themselves very largely to short-term lending and refrain from locking up their money in financing the fixed capital requirements of industry. The banks and their customers remain independent of one another; the banks do not seek control, nor does industry invite it. It is only when bank advances have become frozen that the banks, in their own interests, are forced to take an active part in the affairs of the borrower. They do not interest themselves in the promotion of new companies, and although they may underwrite new issues (i.e., guarantee to take up the stock if the public does not subscribe), they do not make themselves responsible for the success of the company as a business venture once it has been successfully floated.

The policy of the British commercial banks is in striking contrast to that of the investment banks which have been of such importance in the finance of continental countries and in the United States. The German banks, for example, made it their object in their early days to assist industry by providing fixed capital for new companies. After nursing these companies for some time, the banks sold their holdings—often to their own customers. They were generally represented on the management of the companies which they floated, and continued to take an interest in them, providing capital when necessary.[1] The resources of the banks were originally drawn mainly from paid-up capital so that it was of much less importance than in Britain that their assets should be liquid.[2] Later, the growing wealth of the country brought a large increase in deposits and at the same time created a demand for short-term credits (discounts and advances). Thus the balance-sheets of the German banks began to approximate more closely to the English model. But the banks continued to undertake investment banking—as they still do—holding a comparatively small proportion of their assets in industrial securities, and *turning over* a much larger total as one issue after another was handled for successive clients.

**(3) No Re-discounting.**—Perhaps the most unique feature of the British banking system is the convention by which the commercial banks, when their reserves are threatened, do not borrow directly from the Bank of England. Instead, they call in their loans to bill-brokers, and the bill-brokers, acting as a kind of go-between, are "forced into the Bank"—that is, they recover the money which they are being asked to repay by borrowing from the Bank of England. Once the banks have cut down other assets, they are

[1] As each industry tended to be financed by a single bank, it was in the interests of the banks to limit competition between their customers. Hence they generally took a leading part in carrying through schemes of amalgamation within the industries which they specialised in financing.

[2] At the end of the nineteenth century, deposits were still only 50 per cent. of the total resources of the big German banks. In the last forty years, however, the proportion has increased rapidly and is now roughly equal to the proportion in Britain.

able to renew their loans to the bill-brokers and the market ceases to be "in the Bank." Recourse to the Bank of England, therefore, besides being indirect, is rare and transient. In other countries, however, where the discounting of bills is undertaken exclusively by the banks themselves (so that there are no intermediaries corresponding to the bill-brokers) the commercial banks may remain for a long time in debt to the central bank. In the United States, for example, it is the regular practice for commercial banks to obtain funds from the Federal Reserve Banks by rediscounting bills, promissory notes and other "ineligible paper." In France it was formerly the custom for the commercial banks to apply to the Bank of France for rediscount when short of cash. But their resentment at the action of the Bank in entering into increased competition with them for business—contrary to the general practice of Central Banks—and their reluctance to disclose their operations to it, have driven them to other expedients such as the holding of a very large cash reserve.

**The Bank of England.**

The Bank of England differs from the commercial banks in being a central bank—that is, one which is responsible for putting into effect the monetary policy of the country. The government may instruct it, for example, to stabilise exchange rates, or prices, or employment, and it is the Bank's duty to carry out these instructions to the best of its ability, without regard to any loss which its shareholders may suffer in consequence. If the Bank comes into conflict with the government over the most desirable line of policy to be pursued, or over the best way of giving effect to that policy, then the government can, if it chooses, override the wishes of the Bank. But such conflicts appear to be very rare. The Bank, although independent of the Treasury, works in close co-operation with it, and there are joint consultations before any decision of importance is taken. The discretion of the Bank has been increasingly limited in the past few years and it seems destined, in time, to become little more than the agent of Treasury policy.

The activities of the Bank can be most readily understood from a survey of its balance sheet. This is divided—and has been since the Bank Charter Act of 1844—into two separate returns, one made by the Issue Department and one by the Banking Department of the Bank.

**The Note Issue.**

The division of the Bank into two separate departments was part of the general scheme for regulating the currency adopted in the 1844 Act. The defenders of this Act, nicknamed the "Currency School," held that banknotes were money while bank deposits were mere substitutes for money. They were anxious, therefore, to isolate notes from other bank liabilities, and pressed for the separation of the Bank of England—the leading bank of issue—into two separate departments, so as to emphasise the distinction. They would have

liked all notes to be backed by gold, but as this would have meant the withdrawal of an enormous volume of notes from circulation, they agreed on an alternative principle—the so-called "fiduciary principle"—which left the existing note issue unchanged but made it impossible to add to it without the acquisition of additional gold backing. All notes beyond a fixed limit—the fiduciary limit—had to be backed pound for pound by gold, so that if the Bank of England lost gold it required to reduce its note issue, while if it gained more gold it could increase its note issue. It was imagined that this principle would make the supply of money behave exactly as if it consisted exclusively of gold and so would prevent any sudden inflation or deflation. But as was pointed out by the opponents of the Act (the "Banking School") it was inflation of credit rather than of currency which was to be feared; and the Act not only did nothing to regulate credit, but actually made the task of regulation more difficult. To control the supply of money by regulating the note issue was like trying to control the supply of currency by regulating the issue of pennies. The circulation of pennies does not by any means keep in step with the note circulation, nor the note circulation with bank deposits, still less with trade activity. By stressing the importance of notes, therefore, the 1844 Act forced on the Bank a quite inappropriate index of monetary policy—one which, since the note circulation lagged behind trade activity, led to belated and excessively violent action by the Bank to curtail or to expand credit.[1]

*An Account for the Week ending Wednesday, August* 23, 1939.

ISSUE DEPARTMENT.

| | £ *millions* | | £ *millions* |
|---|---|---|---|
| Notes issued— | | Government debt .. | 11·0 |
| In circulation .. .. | 508·1 | Other govt. securities .. | 284·2 |
| In Banking Department | 38·4 | Other securities .. .. | 4·1 |
| | | Silver coin .. .. | 0·7 |
| | | Amount of fiduciary issue | 300·0 |
| | | Gold coin and bullion[2] .. | 246·4 |
| | £546·4 | | £546·4 |

[1] See also page 365 n. Although the separation of the Issue and Banking Departments of the Bank of England was a retrograde step, other provisions of the 1844 Act were of real value. Any possible expansion in the note-issue of the English country banks was arrested by confining the right of issue to banks which already had notes in circulation and by fixing a maximum issue for each. Under another provision of the Act, banks with more than six partners (for example, as a result of amalgamation or of conversion into a limited company) lost the right to issue notes, while the Bank of England was empowered to add two-thirds of the lapsed issue to its fiduciary issue. This ensured a gradual transfer of surviving rights of issue to the Bank of England—a process completed in 1921 by the amalgamation of Messrs. Fox, Fowler & Co. with Lloyds.

[2] At 148s. 6d. per fine oz.

BANKING DEPARTMENT.

| | £ millions | | £ millions |
|---|---|---|---|
| Proprietors' capital .. | 14·6 | Government securities .. | 99·7 |
| Rest .. .. .. | 3·6 | Other securities— | |
| Public deposits .. .. | 22·4 | Discount and advances | 5·7 |
| Other deposits— | | Securities .. .. | 24·3 |
| Bankers .. .. .. | 92·1 | Notes .. .. .. | 38·4 |
| Other accounts .. .. | 36·2 | Gold and silver coin .. | 0·8 |
| | £168·9 | | £168·9 |

The operation of the fiduciary principle is illustrated in the specimen balance sheet of the Issue Department which is printed above.

The liabilities are divided into notes issued to the public (including the joint-stock banks) and notes held in the Banking Department as a cash reserve against the deposit liabilities of the Bank. The backing for the note issue, which is listed on the assets side, consists of the entire gold stock in the possession of the Bank (apart from a negligible sum in the Banking Department), and securities (chiefly Treasury Bills) up to the fiduciary limit of £300 millions.[1] Of the four main items, one—the reserve of notes held in the Banking Department—is purely residual and is governed by the magnitude of the other three. Until recently this meant that it was governed chiefly by the gold holdings of the Bank, increasing as the gold stock was added to and decreasing as gold was withdrawn; for, since the fiduciary limit was fixed and the note circulation comparatively inelastic, the first could cause no change and the second no large or sudden change in the reserve of notes. Now the reserve of notes—again until recently—was the pivot of the British monetary system. The greater the reserve, the more credit the Bank of England could build on it; the smaller the reserve, the greater the pressure on the Bank of England to initiate a general contraction of credit. The monetary policy of the country, therefore, was to a large extent dictated by the size of the Bank's gold stock. The size of the note circulation also played a part. When the note circulation expanded—at Bank Holidays and at Christmas, for example—the reserve of notes automatically diminished, and the drain of notes into circulation tended to bring about a credit restriction exactly like a drain of gold to foreign countries. There is something to be said for restricting credit when gold is flowing out of the Bank to foreign countries, but no intelligible reason can be given why a seasonal drain of notes into circulation should be allowed to influence monetary policy. This was recognised when, in 1937, the practice was introduced of raising the fiduciary limit during the Christmas season and so offsetting the seasonal expansion of the note issue.

[1] The "Government Debt" of £11 millions consists of loans made from time to time to the government between 1694 and 1844 and never repaid. A small amount of silver coin is also held as backing for the fiduciary issue.

Originally the fiduciary limit was absolutely rigid, apart from such increases as took place from time to time when the note-issuing powers of the country banks were transferred to the Bank of England. In the seventy years between 1844 and 1914—a period of rapidly-increasing population and trade—the fiduciary issue of the Bank of England rose only from £14 millions to £18·4 millions, and since the Bank could add only two-thirds of any lapsed issue to its note issue, the *total* fiduciary issue for all banks was diminishing, not increasing. The growing currency requirements of the country—fortunately kept within limits by the development of the cheque system—had thus to be met entirely through the use of more gold sovereigns. Thanks to the discovery of large, untouched gold deposits in three different continents during the period, the necessary gold was obtained without much difficulty. But it was through good luck, not good management, that we avoided drastic deflation.

In 1914 the Bank of England was still prohibited from issuing notes of less than £5 in value. Under the stress of war, however, the government at last decided to issue £1 and 10s. notes. These notes, which were issued by the Treasury, not the Bank of England, were intended to facilitate the withdrawal of gold from circulation into a centralised reserve, and to provide for the increased currency requirements of war-time. After reaching a maximum circulation of £340 millions in 1920, the Treasury note issue was gradually reduced and was finally amalgamated with the Bank of England's issue in 1928. Under the Currency and Bank Notes Act of that year, the fiduciary limit was raised from £19½ millions to £260 millions, and the Bank was allowed to issue £1 and 10s. in replacement of Treasury notes. Some elasticity was introduced into the fiduciary principle by a clause authorising an increase in the fiduciary issue by Treasury Minute for periods not exceeding six months and renewable without Parliamentary sanction up to a maximum of two years. This clause was made use of in August, 1931—just before Britain left the gold standard—when the fiduciary issue was raised for a few months to £275 millions; and again in 1936 when the fiduciary issue was reduced to £200 millions. After a temporary increase to £400 millions in January, 1939, the fiduciary issue was fixed by the Currency and Bank Notes Act of the same year at £300 millions.

If the first Currency and Bank Notes Act stretched the fiduciary principle, the second turned it almost upside down. The essence of the principle is that the credit base of the country—the reserve of notes in the Banking Department—should vary with the gold stock of the Bank of England. But what is now happening is that the gold stock is being adjusted to the reserve of notes. If the Bank finds that its reserve of notes is falling, it applies for gold to a new institution, the Exchange Equalisation Account. Alternatively, it may ask for the fiduciary issue to be increased. On the other hand, if the Bank has a very large reserve of notes, the fiduciary issue may be reduced. Suppose, for example, that the Exchange Equalisation Account has more gold that it wishes and sells some to the Bank of England Then

if the fiduciary issue is reduced by the amount of gold transferred to the Bank, the reserve of notes remains unchanged. The traditional link between gold and the reserve of notes has been broken.

Yet the fiduciary issue continues to exist; some notes are still nominally backed by gold and some by securities. There is no reason why this should be so. There is, in fact, no reason why the Bank of England should continue to hold gold at all. The entire gold stock of the country should be under the control of the Exchange Equalisation Account, and the entire note issue should be backed by securities. Gold should be used only to support the exchange rate; which is the use to which the Account puts it. It has not, and should not have, any more connection with the note issue than it has with silver currency. The only valid reason for maintaining the form of the fiduciary issue when the substance has gone is fear of inflation; the Bank of England has still to apply for gold to the Exchange Equalisation Account (a committee of the Treasury) when its reserve of notes is reduced by inflation. But if the Bank of England is already pursuing a policy sanctioned by the Treasury—and if it is not, then it most certainly should be—the Treasury cannot refuse its assistance in the carrying out of that policy. So long as the Bank of England and the Treasury are in agreement it does not matter much whether gold is stored in the name of the one or of the other. But it does make for clear thinking when gold is set aside explicitly for the purpose for which it is really required.[1]

**The Banking Department.**

The operations of the Banking Department can be more briefly described. The share-capital of the Bank, which was fixed at its present figure in 1844, is owned by members of the public. Reserves and undistributed profits are put under the curious heading "Rest." Deposits, the chief liabilities of the Bank, are grouped under three headings corresponding to the three functions of the Bank as banker to the government, to the commercial banks, and to other customers.

**Public Deposits.**

As the financial agent to the government, the Bank receives taxes, pays interest on the National Debt, and supervises the issue of all government loans. The government requires, therefore, to keep large accounts with the Bank—for example, the Exchequer Account, and the account of the Commissioner of the Public Debt—and these accounts are returned as "Public Deposits." Very large fluctuations may take place in "Public Deposits" at times when tax payments are due or when disbursements of interest have to be made, and these fluctuations, if not offset, are liable to exercise a disturbing influence on the money market.

[1] Since this was written, the Bank of England's gold reserve has been transferred to the Exchange Equalisation Account (in September 1939), and the fiduciary issue has been increased to £1,000 million (in 1943).

**Bankers' Deposits.**

This item includes only deposits made by the British Commercial banks. Since these deposits, together with cash in hand, form the first-line reserves of the commercial banks, any change in their amount is generally associated with a change in credit policy. An increase enables the banks to lend more freely and a decrease forces them to restrict credit.[1]

**Other Accounts.**

Although the Bank has withdrawn from competition with the commercial banks, it still carries on an ordinary banking business. Its customers include foreign central banks, colonial banks and governments, financial houses in the City, and a number of private customers. The accounts of these customers are included under "Other Accounts."

Turning to the assets side of the balance sheet we find first :

**Government Securities.**

These include British government bonds and Treasury bills purchased by the Bank on its own account, and temporary advances to the Exchequer (called "ways and means advances"). When the Bank tries to increase or reduce its holdings of government securities by buying or selling them, its operations are described as "open market operations," since they are undertaken on the Bank's initiative in the open market. Such open market operations, as we shall see, are part of the Bank's technique for controlling credit.

**Other Securities.**

Under the sub-heading "securities" are included securities other than those issued by the British government ; for example, the bonds of colonial and foreign governments, the debentures of industrial companies, and commercial bills purchased by the Bank on its own account. The sub-heading "discounts and advances" is made up partly of bills discounted at the request of the Bank's customers and of advances to these customers, and partly of advances to the money market. The minimum rate at which the Bank will discount first-class bills is known as Bank Rate ; but customers of the Bank may have their bills discounted at the ordinary market rate of discount which is almost invariably below Bank Rate. The reason for keeping Bank Rate above market rate is to prevent active competition between the Bank of England and the money market. The Bank holds aloof as "the lender of last resort," discounting bills for its own customers, but not bringing pressure to bear on the open market by the purchase or sale of bills except when it wishes to make discount rates fall or rise. By this policy, the Bank enormously reinforces its control over the money market since its intervention is not resented as that of a competitor and takes all the more effect by being concentrated at the margin.

[1] See Chapter 25.

Since discounts and advances to customers are relatively small, the total for " discounts and advances " is a good index of the extent to which the market is " in the Bank." Normally, as we have seen, the market does not borrow from the Bank. But when there is a shortage of funds, and bill brokers are deprived of the call money lent to them by the joint-stock banks, they are forced to borrow from the Bank of England. Their borrowings, which are made at a rate of 1–2 per cent. over Bank Rate, are reflected in an increase in " discounts and advances."

Notes.

This item has already been explained. For all practical purposes it consists of " free gold "—gold surplus to the legal requirements governing the fiduciary issue. The Bank might just as easily hold this gold in the Banking Department as issue notes against the gold and hold these notes instead. In practice, however, it adopts the second course.

The reserve of notes, together with a small amount of gold and silver coin, forms the Reserve of the Bank, corresponding to the cash reserve of the joint-stock banks. The ratio between the Reserve and the Bank's deposits is called the Proportion, and this, though more variable than the reserve-ratio of the joint-stock banks, is kept within fairly narrow limits. Up to 1914 the Proportion was rarely allowed to fall below 40 per cent.; since then it has been as low as 14 per cent.; now it is kept in the neighbourhood of 30 per cent.

The Proportion is the traditional barometer of British monetary policy. If it falls, a restriction of credit is likely to be in store; if it rises, we can look forward to a period of easy money. Why this is so—or rather, why it used to be so, for the barometer can now be rigged to give any desired forecast—we shall discuss in the next chapter.

## CHAPTER 25

## THE SUPPLY OF MONEY[1]

### Banks and Savings

THE primary function of banks is to mobilise savings. They gather driblets of spare cash, much as the buckets of a water-wheel gather driblets of water, and harness them to the tasks of enterprise. The odds and ends of savings which we lodge with the banks are formed into a large stock of capital which can be lent at low rates of interest in amounts commercially useful. Our individual deposits are liable to be withdrawn at any moment; but in the aggregate they are relatively permanent, and a proportion consistent with

[1] Throughout this chapter the volume of bank credit is assumed to be constant.

experience and prudence can be lent with safety for definite periods of varying length. Thus the banks borrow in small amounts and lend in large amounts; they borrow short and lend long; above all, they borrow liquid capital and lend it in much less liquid forms. Each transformation enhances the productivity of loan capital; for the larger the loan, and the longer and more definite the period for which it is made, the more valuable, in general, are the uses to which it can be put. The banks, therefore, in mobilising capital, add to its productivity and render a service of great advantage to the community.

### Banks and Credit

The function of mobilising savings has become entangled with a second, and quite distinct, function—the supply of the means of payment. If the only money in use was gold, the supply of it would be governed exclusively by the activities of the gold-miners. The savings which we lodged with the Bank, and which they in turn lent, would consist of gold coins; and although the introduction of banking would make money circulate more rapidly it could not add to the supply. When banknotes came into use, however, the banks acquired power to *manufacture* money. By increasing their loans they could put more notes into circulation, since the loan was withdrawn and paid away in notes, and these notes, being generally acceptable, were to all intents and purposes, money. Similarly they could, by calling in their loans (that is, by "restricting credit"), reduce the supply of money. The powers of the banks were limited only by the reserve of gold which they thought fit to keep. If, for example, they held rigorously to a reserve ratio of 20 per cent. (i.e., if they held £20 in gold against every £100 in notes), they would be able to issue notes only to the value of five times the gold deposited with them. Whenever an additional deposit was made, the banks would be able to lend five times the value of the new deposit, and so add four times its value to the amount of money in circulation. A withdrawal of deposits would react automatically to reduce the note-issue in the same proportion by forcing the banks to call for repayment of their loans.[1]

**Creation of Credit.**

When gold disappears from circulation, and the issue of notes is limited or taken over by the government, the power of the banks to create money does not by any means come to an end. For the banks control the supply of bank deposits much as they formerly controlled the supply of banknotes and it is deposits, not notes, which are nowadays the chief medium of exchange. Just as they were once able to vary their note issue by increasing or diminishing their loans, so the banks can now vary their deposits by exactly the same tech-

[1] Even under a regime of gold only, the banks would add to, or diminish their loans whenever gold was deposited or withdrawn. But they would be able to lend only a proportion of any new deposits, not several times their value.

nique. By lending more freely they can *create* deposits, and by restricting credit they can extinguish them. The mechanism by which they do this calls for detailed consideration.[1]

Let us begin by supposing, for the sake of simplicity, that there is only one bank and that it keeps a constant reserve ratio between cash and deposits of 10 per cent. Then if, for any reason, the public leaves an additional £1 million on deposit with the bank, it will be necessary to hold against this deposit a reserve of £100,000. This leaves the remaining £900,000 at the disposal of the bank for making loans to customers or for investment in bills, government securities, etc. If the £900,000 is lent, it may be withdrawn and paid away in wages or for goods. But it is not necessarily withdrawn. The borrowers may pay for the goods which they buy with a cheque drawn on the bank; they may even pay their workmen (as in America) by cheque. These cheques will be paid into the bank—*re-deposited* with it—and the money will be credited to the accounts of the recipients. No cash will have been withdrawn, so that the bank will still have its £900,000. At the same time the bank's deposits will have been increased by the value of the cheques paid in; that is, by £900,000. At this stage, therefore, the total increase in deposits is £1·9 million—£1 million lodged on the initiative of the public, and £900,000 created by the banks themselves. But there is no reason why the process of credit creation should end here. The bank need only set aside an additional £90,000 as a reserve against the new deposits and lend the remaining £810,000. This in turn will be re-deposited when cheques are drawn against the loans, the bank will set aside an appropriate reserve and the balance will be lent. So long as no cash is withdrawn by the public there is only one possible ending to this process; the lending of ten times the original deposit of £1 million. This will absorb the whole of that sum into the reserves of the bank and restore the normal ratio between reserves and deposits.

Presumably, however, borrowers will want to withdraw in cash some part of the advances made to them. Wage-payments, for example, are generally made in notes, not by cheque. At some stage, therefore, the banks will be asked to provide notes and will only be able to do so by drawing on their reserves. If all of these notes remained indefinitely in circulation, the power of the banks to create credit on the base afforded by the original deposit would be materially reduced. But most of the notes will return to the banks as additional deposits whenever they come into the possession of people who keep banking accounts (e.g. shopkeepers). Only if the public finds it necessary to hold more notes in order to transact its business is there a drain of

[1] The fact that loans create deposits serves to remove some of the ambiguity lurking in the phrase "creation of credit." For credit is sometimes used in opposition to "currency" to mean "bank deposits"; and sometimes to mean a loan (by a banker, shopkeeper, or anyone else). If bank loans create deposits, it does not matter very much whether we have in mind the additional lending or the resulting increase in deposits when we say that "credit" has been created.

notes from bank tills to swell the normal quantity in circulation. In view of the expansion in incomes and in business activity to which credit creation leads, we are safe in concluding that there will be *some* expansion in the note circulation, and some check, therefore, to the power of the banks to add to their loans. But the ultimate net increase in notes in circulation will be much less than the initial withdrawals.

The position is more complicated, but the final result is unaltered when there are several banks competing for the deposits of the public. Suppose that the Midland receives deposits of £1 million, and lends the usual nine-tenths. Even if the borrowers withdraw nothing in cash, paying away the full value of their loans by cheque, the Midland will nevertheless begin to lose some of its cash. For those to whom the cheques are paid may be customers of other banks, and when they pay in their cheques, will create deposits in *these* banks. The Midland, therefore, will be faced with a debit balance at the Clearing House, and will require to surrender cash to its competitors to the value of the cheques drawn on it. If, for example, there are nine banks, all equally large, it is probable that only £100,000 out of the £900,000 lent by the Midland, will be redeposited with it, while each of the other banks will acquire deposits amounting to £100,000. The Midland will then be able to lend only an additional £90,000 out of the deposits which it has created, and if eight-ninths of these loans again find their way into deposits with competitors, it will be left with only £10,000 out of the next batch of deposits. By successive loans, the Midland will be able to create deposits with itself to the aggregate value of £1 million making, with the original deposits of £1 million, a total increase of £2 million. This is far below the credit expansion which was possible when there was one bank only. But if the other banks, which between them must have acquired £800,000 in cash, begin lending in *their* turn, the final result will be exactly as before. So long as the banks lend in step with one another and try to keep their reserve ratios constant, it makes no ultimate difference whether there is only one bank or several.

So far we have assumed that the initiative is taken by the public in adding to its deposits. But in fact it is rarely the public which sets the process of credit creation in motion. The cash requirements of the public are inelastic and vary within fairly narrow limits, more in response to changes in credit than as their cause.[1] It is the banks themselves which generally initiate a credit expansion. They are free at any time to increase their loans without waiting for the public to bring in notes and deposit them. The proceeds of bank loans will sooner or later be lodged as deposits; instead of deposits giving rise

[1] Once an expansion in credit has increased our incomes or raised prices, we generally keep more cash by us (as has been argued above), while a fall in incomes and prices leads us to keep less by us. So far as they go, therefore, changes in our holdings of cash *stabilise* credit creation by eating into bank reserves in the upswing and adding to them in the downswing: they do not *initiate* or *aggravate* credit creation except in special circumstances.

to loans, the loans will create deposits. If any other bank asset is increased, the effect is the same. A purchase of government securities, for example, is paid for by cheque, and when the cheque is paid into the bank the seller's account is credited with the sum due. The bank obtains the securities and the seller receives a bank deposit. If, therefore, the banks wish to increase or diminish total deposits they need only buy or sell securities.

**The Joint-Stock Banks and Credit Creation.**

Although the banks have power to create credit, it is conceivable that they might never make use of this power. Total bank deposits are the product of the absolute amount of the bank's cash reserves and the ratio which the banks seek to maintain between reserves and deposits. But no bank can, of itself, add to its cash reserves except by drawing on some more profitable asset; and if the banks have already lent " up to the hilt " (i.e. until their reserve-ratio is as low as is consistent with sound banking) they will not want to reduce their reserve ratio by further credit creation. It might seem, therefore, as if the banks had power only to restrict credit, not to create it. In fact, however, British banks have not lent up to the hilt. That part of their reserves which they hold with the Bank of England is almost certainly in excess of their maximum requirements. If they choose to expand credit and reduce their normal reserve ratio, therefore, they can do so. Indeed they have done so. Between 1926 and 1929 deposits with the London clearing banks rose from £1,600 million to £1,732 millions, while in the meantime their reserves *fell*.[1] Nevertheless, independent action by the commercial banks is as rarely the motive power behind credit creation and contraction as independent action by the public. The real power rests with the Bank of England. For it is the Bank of England which determines the absolute amount of the commercial bank's reserves and it is these reserves which, given constant reserve ratios, control the volume of credit.

**The Bank of England and Credit Creation.**

The Bank of England's position in relation to the commercial banks is parallel to their relation to the public. The public holds money on deposit with the commercial banks which in turn hold money on deposit with the Bank of England. Just as the commercial banks can create deposits with themselves, so also can the Bank of England. The Bank, it is true, does not lend directly to the joint-stock banks as they do to the public. But it can, by " open market operations," achieve a similar result. If the Bank wishes to pursue an expansionist policy, it buys government securities paying with a cheque drawn on itself. This cheque is paid into one of the commercial banks, which credits it to the seller. The bank then collects the proceeds from the Bank of England by having them added to its balance there. It has received additional deposits from the public—deposits originally

[1] Report of Committee on Finance and Industry (Macmillan Report), page 287. The lead was taken by the Midland Bank.

created by the Bank of England—and it is, therefore, in a favourable position to start a cumulative expansion of credit of the type which we have already analysed. Similarly if the Bank of England sells government securities, it reduces bankers' balances with itself and forces them to contract credit. Bankers' balances with the Bank of England are the pivot of the British monetary system, and it is through its control over their amount that the Bank of England is able to regulate the creation of credit.

It is important to notice upon how slender a foundation this power of the Bank of England rests. Without control over the bankers' reserves which are centralised in its keeping, the Bank would be unable to dominate monetary policy. Yet there is no law which forces the joint-stock banks to keep balances with the Bank, and no interest is paid on them. It is simply the convenience and prestige which come from leaving some minimum reserve with the Central Bank that prevents the joint-stock banks from holding their reserves in some other form and from varying them in a manner which might obstruct the declared policy of the Bank and of the government. There can be little doubt, however, that if such obstruction did take place—if, for example, the joint-stock banks attempted to withdraw their reserves from deposit with the Bank of England and to place them abroad with foreign banks—a legal minimum would soon be set to the proportion of their deposits which the banks hold at the Bank of England.[1]

**Limits to the Powers of the Bank of England.**

We have seen that although the joint-stock banks have power to create credit, that power is so limited by the need to keep a fairly constant reserve-ratio that ultimate control passes to the bank which holds their reserves and determines their aggregate amount—the Bank of England. We have now to see in what ways the discretion of the Bank itself is fettered by customary or legal limitations.

Originally, before it developed into a bankers' bank, the Bank of England held a reserve of legal tender from the same motives as other banks. It had to be ready, like them, to redeem its notes, or to repay its deposits, in legal tender, and its policy was guided, like theirs, by the reserve against these liabilities which it considered to be adequate. It neither had, nor believed itself to have, any special powers over the money supply that distinguished it from the innumerable private banks with which it competed. The supply of money was governed, first by the stock of legal tender money in the country (consisting at that time of gold and silver coins); second, by the amounts of it which the public deposited with the banks; and third, by the reserve-ratios which the banks saw fit to maintain. Given the customary practices of the banks, and the currency requirements of the public, the controlling factor was the supply of legal tender. Any addition to the gold stock either passed directly into circulation as currency or

[1] Such a minimum is enforced in the United States on all banks which are members of the Federal Reserve System.

was deposited with the banks, increasing their reserves and providing a base on which a larger structure of credit could be built. A drain of gold had precisely opposite effects, depleting the reserves of the banks and initiating a general deflation.

In course of time, the Bank of England became the custodian of the country's central gold reserve—the natural result of the Bank's long monopoly of joint-stock banking and the consequent absence of serious competition from banks of equal strength. The flow of gold into or out of the country was canalised through the bottle-neck of the Bank's gold stock, so that it was this stock which bore the impact of any change in the demand for legal tender. The commercial banks, holding their reserves either at the Bank or in Bank of England notes, were content to leave to it the responsibility of ensuring that there should always be a sufficiency of gold. It was a responsibility which the Bank was slow to recognise; but, recognised or not, it gave to the monetary policy of the Bank of England a place of enormous importance in the regulation of the money-supply.

So long as Britain remained on the Gold Standard, this concentration of the reserves of the commercial banks under the control of the Bank of England left the Bank with only limited powers over the creation of credit. The Bank was obliged, in conformity with the Gold Standard, to convert its notes freely into gold at a fixed price—before 1914 into gold sovereigns. Its over-riding care, therefore, was to hold a gold reserve large enough to cope with the maximum call that might be made upon it.[1] Its increased responsibilities made the need to hold an adequate gold reserve greater, not less, than before, and it was this need, rather than any consideration of its own, or of the public, interest that determined its ability to create credit. Its powers, like those of other banks, were limited by the reserves which it held and by the rigidity of the reserve-ratio which it sought to maintain. The Bank was, in fact, in the same position in relation to other central banks as one commercial bank is in relation to its competitors. Just as commercial banks hold reserves in order to meet debit balances against them at the Clearing House and settle their differences by transfers of cash, so central banks, under Gold Standard conditions, hold reserves to meet debit balances of payments between countries, and settle unfavourable balances by transfers of gold. For the most serious threat to their reserves—and the chief reason, therefore, for holding them—is the danger of a drain of gold to foreign countries and into foreign banks whenever the balance of payments is unfavourable (that is, whenever the payments which have to be made to foreign countries are in excess of the payments due from them).[2] Such a drain of gold is parallel to the loss of cash by one commercial

[1] This gold reserve is represented by the reserve of notes in the Banking Department. (See page 321).

[2] The export of gold from the reserves of the Central Bank meets the deficit in the balance of payments until action can be taken to restore equilibrium. (See page 362 *et seq.*).

bank to another. Just as the transfer of cash forces one bank to contract, and allows the other to expand credit, so a drain of gold from a central bank tends to start a credit contraction and an increase in gold stocks to facilitate credit expansion. But whereas the commercial banks hold to a more or less fixed reserve-ratio, central banks allow theirs to vary widely. Their policy is guided by the state of their reserves but not dictated by it.

**The Money-Supply under Gold Standard Conditions.**

Broadly speaking, therefore, the money-supply in Britain was governed, while we were on the Gold Standard, by three main factors: the supply of gold, the customary practices of the public and of the banks, and the monetary policy pursued by the Bank of England.

**(1) The Supply of Gold.**—Any addition to the supply of gold—for example, through an increase in the output from the mines—sooner or later reacted on the Bank of England's gold stock and so on the reserve of notes. This broadened the base of credit and, by increasing the Proportion, caused the Bank to make open market purchases of securities. These in turn increased bankers' deposits and encouraged the joint-stock banks to create additional credit. Periods following important gold discoveries, therefore, have generally been accompanied by an expanding money-supply.

**(2) Customary Practices.**—If the banks maintain fixed proportions between their reserves and their deposits, and if the public keeps a fixed proportion between cash in hand and money at the bank, then the total supply of money will be a fixed multiple of the Bank of England's reserve. But any change in these proportions will react on the money-supply. A reduction in the note circulation, for example, will release gold held as backing and increase the "free gold" reserve of the Bank of England, with the same results as an increase in the total gold stock.

**(3) Monetary Policy.**—No bank keeps its reserve-ratio absolutely fixed and the Bank of England in particular allows the Proportion to vary whenever the monetary policy which it is pursuing makes such variations unavoidable. If, for example, the Bank wishes to refrain from deflation even when it is losing gold, it will allow the Proportion to fall below the normal figure. But there are clear limits within which the Bank can permit such variations. It cannot watch its reserve melt away without taking action sooner or later and if it waits too long, it may be forced to take far more drastic action than would have been necessary earlier.

**The Money-Supply off the Gold Standard.**

When a country leaves the Gold Standard—that is, when its money ceases to be convertible into gold on stated terms—the money-supply need no longer be linked to gold. The Central Bank may continue to hold a gold reserve and use it for much the same purposes as before.

But it does so—unless some *other* legal restrictions survive[1]—as a matter of policy, not in conformity with the provisions of its charter. The Bank and the government are free to manage the money-supply as they think fit; their discretion is not limited by any arbitrary rule such as convertibility into gold but only by such alternative rules as they may devise in formulating a new monetary policy.

## CHAPTER 26

## THE VALUE OF MONEY

### The Measurement of Changes in the Value of Money

PRICES fluctuate. Not simply individual prices—the price of potatoes, for example—but something which we think of vaguely as the general level of prices or as the cost of living. A general rise in prices is most obvious in war-time when the cost of living goes up sharply; we find that our money does not go so far in buying goods—that its purchasing power or value has fallen because the things which we are in the habit of buying have become dearer. A general fall in prices often takes place at the end of a war or in a trade slump; the cost of living is reduced and the value of money increases correspondingly. Although these fluctuations in prices and in the value of money have been familiar to us for a very long time, it was not until the middle of last century that any serious attempt to *measure* them was made. People were content to rely on personal impressions or to point to particular changes in prices as indications of the general trend. Then Jevons hit on the idea of trying to measure the *average* change in prices by means of index numbers.

**Index-Numbers.**

The first step in constructing an index number of the cost of living is to obtain quotations of the market prices of a representative selection of commodities in the years which we wish to compare. We then choose some year as a basis of comparison and relate prices in every other year to prices in this base year. Suppose, for example, that the price of a loaf is 8d., 6d., 7d. and 9d. in Years I, II, III and IV, and that we have decided to use Year I as our base year. Then we represent the price of a loaf in Year I by the number 100, and the price in Years II, III and IV can then be represented by the numbers 75, 87·5 and 112·5. These numbers are known as price-relatives. The prices of other commodities can be dealt with in the same way, each being expressed in terms of price-relatives. The convenience of these price-relatives is obvious; they allow us to

[1] For example, those governing the note issue. After 1931, notes had still to be backed by gold although not convertible into gold

read off at once the percentage change in price since the base year so that we know exactly in what proportion the price has risen or fallen. Once we have a complete series of price-relatives for each commodity, we can calculate the average change in prices by adding together the price-relatives for each year and dividing by the number of commodities.[1] If we were dealing with only five commodities—say, bread, butter, milk, eggs and potatoes—we might construct a table like this :—

TABLE A.

PRICE-RELATIVES.

| | Year I | Year II | Year III | Year IV |
|---|---|---|---|---|
| Bread .. .. | 100 | 75 | 87·5 | 112·5 |
| Butter .. .. | 100 | 90 | 95 | 105 |
| Milk .. .. | 100 | 105 | 110 | 120 |
| Eggs .. .. | 100 | 110 | 80 | 90 |
| Potatoes .. .. | 100 | 70 | 200 | 120 |
| Total.. .. | 500 | 450 | 572·5 | 547·5 |
| Average .. | 100 | 90 | 114·5 | 109·5 |

The numbers 100, 90, 114·5 and 109·5 are index-numbers of the cost of living in Years I, II, III and IV, and they tell us in what proportion prices have risen or fallen since the base year (Year I). In Year II prices were 10 per cent. lower, in Year III 14·5 per cent. higher, and in Year IV 9·5 per cent. higher than in Year I. A corresponding but opposite change has taken place in the value of money; in Year II, for example, it had $\frac{100}{90}$ or 111 per cent. of its purchasing power in Year I.

This was the procedure by which the earliest index-numbers were calculated. It was not, however, altogether satisfactory. The most obvious objection was that it attached equal importance to each commodity included; equal weight was given to a large change in the price of bread and a large change in the price of chocolate. But it is obvious that even a comparatively small change in the price of bread affects the cost of living much more than a very large change in the price of chocolate. We must find some way, therefore, of giving due weight to each change in price when we are constructing index numbers.

[1] Sometimes the price-relatives instead of being added are multiplied together. If there are $n$ commodities, the $nth$ root of the product is then taken. This gives us the Geometric Mean as opposed to the Arithmetic Mean, or simple average, obtained by adding the price-relatives together and dividing by $n$.

This is done by taking a " weighted " instead of a simple average of price-relatives. Suppose, for example, that the relative importance of the five commodities in Table B can be represented by the numbers 10, 3, 4, 1 and 2 respectively. Then these numbers are called " weights." They are multiplied by the appropriate price-relatives, and an average is struck by dividing the total for each year by the sum of the weights instead of, as in Table A, by the number of commodities.

TABLE B.

| | Price-Relatives | | Weight | Expenditure-Relatives (Price-Relatives x Weight) | |
|---|---|---|---|---|---|
| | Year I | Year II | | Year I | Year II |
| Bread .. | 100 | 75 | 10 | 1,000 | 750 |
| Butter .. | 100 | 90 | 3 | 300 | 270 |
| Milk .. | 100 | 105 | 4 | 400 | 420 |
| Eggs .. | 100 | 110 | 1 | 100 | 110 |
| Potatoes | 100 | 70 | 2 | 200 | 140 |
| Total .. | 500 | 450 | 20 | 2,000 | 1,690 |
| Average | 100 | 90 | — | 100 | 84·5 |

The result of this procedure, as is illustrated in Table B, is to show a fall in prices between Year I and Year II of 15·5 per cent. compared with a fall of only 10 per cent. by the earlier method. The weighted index number has fallen further than the unweighted, because the heaviest fall in prices has taken place in the commodities entering most largely into consumption (e.g., bread). It will be observed that the last two columns of Table B are headed Expenditure-Relatives. The reason for this is that if the figures in column 4 represent our expenditure on each commodity in Year I, our bill for the same commodities at the prices ruling in Year II is exactly represented by the figures in column 5. The weights, it will be remembered, correspond to the relative importance of each commodity; in other words, to the proportion between our expenditure on each commodity and our total expenditure on all commodities taken together. If we give bread as large a weight as all the remaining commodities taken together, this *means* that half of our total expenditure goes on bread; of 2,000 units of purchasing power we expend 1,000 on bread. A fall in the price of bread by 25 per cent. will cut our bread bill by 25 per cent. We require only 750 units of purchasing power to buy what previously cost 1,000 units. Similarly with butter, milk, eggs and potatoes. At the level of prices ruling in Year II we can make the same purchases with 1,690 units of purchasing power as we made in Year I with 2,000 units. The purchasing power of money has increased in the ratio 2,000 : 1,690, and the cost of living

has fallen in the ratio 1,690 : 2,000. What we are really doing, therefore, when we weight an index number of the cost of living by the method used in Table 00 is to calculate the cost in successive years of a constant bundle of commodities—the bundle on which a typical family may be supposed to spend its income. If prices fall by 10 per cent. then this bundle will cost 10 per cent. less ; if prices rise by 10 per cent., then it will cost 10 per cent. more to maintain the old standard of living. By watching how the cost changes from year to year, we can measure changes in the price-level and in the value of money.

But how do we decide which commodities to include in the standard bundle and how much of each to include ? How, in other words, do we arrive at the proper *weights* ? The answer is : from family budgets. We require to gather from a large number of families particulars of how they spend their income. We can then calculate the proportion of their income which people spend *on the average* on each commodity. This information, together with price-quotations from different parts of the country for each commodity, is all that we need in order to construct an index number of the cost of living.

The method of constructing index numbers which is outlined above is the one used by the Ministry of Labour in its official index—the only cost of living index which is available for Great Britain. The Ministry's index, which is intended to apply only to changes in the cost of living of manual workers, is based upon an investigation made in 1904 into urban working class budgets. As a result of this investigation, the following weights were adopted :—

| | | | |
|---|---|---|---|
| Food .. .. .. | 60 | Fuel and light .. .. | 8 |
| Rent and rates .. .. | 16 | Miscellaneous .. .. | 4 |
| Clothing .. .. .. | 12 | Total .. .. .. | 100 |

A new investigation was undertaken in 1938 and the weights will be revised when the results of this investigation are published. The price quotations used in the construction of the index are obtained weekly from 630 towns and villages and these quotations are averaged for each commodity before the price-relatives are calculated.

**Difficulties in the Construction of Index Numbers.**

If all prices moved together, there would be no need to construct index numbers ; the change in any one price would measure the change in the average level of prices. But prices do not move together. Some rise more than others in a general upward movement ; some may even fall while the others are rising. The core of our problem, therefore, is to average one change against another. But—and here we come to the first, and major, difficulty—everyone is not affected in the same way by any given change in price. A rise in the price of Rolls-Royce cars, for example, means a higher cost of living to one group of people but leaves others unaffected. A rise in the price of bread raises the cost of living far more to a poor man than to a rich man. Different people spend money on different things and buy

different amounts of the same things. Thus the weights appropriate to one man's expenditure may be quite inappropriate to another man's expenditure. For no two persons or families will the weights be exactly the same; strictly speaking, therefore, we ought to calculate a separate index number of the cost of living for everyone, since, because of differences in taste and in income, in size of family and in place of residence, the bundle of goods which each of us consumes differs from the standard bundle consumed by that compromise—the average family. In practice, it is found to be unnecessary to construct a large number of index numbers; even with quite different weights from those used at present much the same results are obtained. It is highly desirable, nevertheless, that we should at least have separate indices for people in different income groups. An index which is intended to measure changes in the working class cost of living may be seriously misleading if it is taken as a measure of changes in the value of middle and upper class incomes. At a higher standard of living, a much higher proportion of income is spent on direct services—the services, for example, of doctors, lawyers, teachers, waiters, gardeners, etc.—and these services are naturally given much less weight in an index intended to apply to working class standards than they would be in an index designed to represent changes in the purchasing power of money to the well-to-do.

A second difficulty is that people buy different things at different times. New tastes and habits of mind develop and we vary our customary purchases. New goods come into use and a place has to be found for them in our budget. One commodity becomes relatively dear and another relatively cheap, and we adapt our consumption to take advantage of the change in relative prices. Some services (e.g., education) are taken over and provided free by the State so that they drop out of our budget; for others (e.g., insurance) we are forced by law to make provision. Thus the bundle of goods which we purchase in Year II differs from the bundle which we purchased in Year I. Exact comparison, therefore, between the purchasing power of money in Years I and II becomes impossible. For we can make such comparisons only when money is used to make the same purchases in both years.[1] Suppose, for example, that we calculate the cost in Year II of the purchases which we made in Year I and then, reversing the procedure, calculate the cost in Year I of the purchases which we

[1] This is not altogether correct. If I know that £300 in Year I would have left me as well off as £400 in Year II then I am justified in saying that (for me) money has lost a quarter of its purchasing power. I am justified in saying so even although my purchases in Year I with £300 might have been quite different from my purchases in Year II with £400. But since the comparison is in part a subjective one I have no means of *proving* to the satisfaction of other people that money has lost a quarter of its purchasing power; and they are at liberty to point out that, while my judgment may be sound given my *present* tastes and values, I might have expressed a different judgment in Year I, with the tastes and values that I had *then*.

made in Year II.[1] There is no guarantee that these two calculations would yield the same result; the change in the cost of living measured *backwards* may be quite different from the change measured *forwards*. Nor is there any reason why we should rule out one calculation as less true than the other; both can claim to measure the "true" change in the purchasing power of money, according as we look backwards or forwards. The wisest course is probably to average the results of both calculations, without speculating too deeply on the precise meaning of the average obtained.

The practical importance of the difficulty raised by changes in purchases is generally small in the short run and increasingly great as time wears on. The main items of expenditure in our budget change little from year to year, and the other items play only a subordinate part in causing changes in the cost of living. The introduction of new commodities is hardly rapid enough to have much effect on the cost of living within a short space of time. Even the consumer's efforts to cut down expenditure on relatively dear commodities and to increase his purchases on relatively cheap ones (e.g., the regular oscillation between butter and margarine as the gap in price widens and narrows) is rarely on a sufficient scale to cause any serious error in our calculations. But when we take a longer view, the problem is by no means a trivial one. He would be a bold man who undertook to compare the value of the £1 in 1066 with its present value.

It might be thought that changes in purchases could be satisfactorily represented by changes in the weights used in the construction of index numbers. We might revise the weights every year in accordance with observed changes in consumption and link prices in each year with prices in the succeeding year. The series of links might then be formed into a chain by letting 100 represent the price level in some base year and attaching each link—each successive change in prices—in turn. Such a chain index number, however, would disregard rather than solve our problem. It would be based on the assumption that changes in purchases in adjacent years were negligible. But it is precisely these changes which we wish to take into account. If we give up using butter when margarine comes on the market, an index of the cost of living ought to reflect the advantage to us of the new and cheaper substitute. But all that a chain index does is to attach more and more importance to changes in the price of margarine as our consumption of it increases. This is an advance on the more usual procedure of assuming that we eat just as much butter as before. But in the long run it is little more satisfactory. The cumulative advantages which we derive from the new commodities introduced year by year are equally disregarded.

There are also a number of practical difficulties in the construction of index numbers. First, we must be careful in our choice of a base

[1] Neither calculation can be exact. Some goods available in Year II were not available in Year I, while other goods, available in Year I, will not be available in Year II.

year. A year should be chosen in which prices are in a fairly normal relationship to one another, and in which our expenditure is distributed between each commodity in a fairly normal way, thus providing a satisfactory base from which to measure off changes in prices and weights which are in keeping with our normal consumption. Sometimes the average of a series of years is chosen as a base, and this has the advantage that the resulting index cannot be biased by the peculiarities of any single year. Secondly, we must exercise great care in making comparisons between different prices, since these prices may be paid for goods or services which do not remain constant from year to year. We may be given goods of better quality; for example, the butter which town workers buy nowadays is greatly superior to the butter on sale half a century ago. Or we may be given better service—as when bottled milk from disease-free herds is delivered on our doorstep. Or we may be given better terms of credit as a result of the growth of banking facilities. Even in the short run what are really changes in price may be disguised as changes in quality or in service; for example, shopkeepers who are in the habit of issuing price lists will find it convenient to charge a constant price and provide goods of better or inferior quality as wholesale prices vary. Thus the extraction of price-relatives from a mass of information about prices is no easy task. Unless we are deliberately setting out to "cook" our index numbers, we must make sure that all price quotations are strictly comparable.

**Other Index-Numbers.**

Money, as we have seen, has a different purchasing power for different people. It has also a different purchasing power for different purposes. We have been assuming up till now that the only purpose for which money is used is the purchase of finished goods, and we have been trying to devise a method of measuring changes in the purchasing power of money to consumers of finished goods. But money is used also to buy other kinds of goods—raw materials, semi-finished goods, machinery, exports, imports and so on—and its purchasing power over these goods may change in a different direction or to a different degree from its purchasing power over finished goods. To look at things the other way round, instead of one price level—the price level of finished goods or the retail price level—there are a large number of price levels corresponding to the different groups of things that we can buy for money. There is the wholesale price level, the export price level, the price level of capital goods, the price level of bonds, the wage level, and a whole host of other price levels. A separate index number can be calculated to measure changes in each price level, and the method of constructing these indices is exactly the same as that used in constructing a cost-of-living index number. The Board of Trade Index of Wholesale Prices, for example, makes use of wholesale price quotations for 200 commodities, each of which is given a weight in accordance with the value of wholesale dealings in it in 1930. Sub-indices have

also been constructed to measure changes in the price levels of "basic materials," "intermediate products," and "manufactured articles."[1]

These different price levels do not move together, and it is for this reason that changes in the purchasing power of money are so important. In a general movement of prices up or down, some prices run ahead of others; wholesale prices, for example, outstrip retail prices, and wages and other costs lag behind both. Normal profit margins and price relationships are upset, and strains and stresses are set up in the price structure. The pressure on productive activity at various points can be gauged by the use of index numbers to measure the spread between price levels. By comparing any two index numbers we can see at once where normal price margins have been disturbed, and on what scale the disturbance has been. By comparing indices of export and import prices, for example, we can see whether we are obtaining imports more cheaply in terms of the exports which we send in payment. By comparing indices of the price of farm products and of manufactured articles, we can see how much more heavily a fall in price bears on farmers than on the rest of the community. Index numbers sum up the changes that re-orientate business activity; they make it immensely easier for us to interpret these changes intelligently and to control them by wise planning.

### The Cause of Changes in the Value of Money.

Since 1914 we have all learned, some of us at considerable cost, that the value of money changes. In Germany the mark lost almost the whole of its purchasing power in a single year (1923); the French franc lost four-fifths of its purchasing power by 1926; and the pound, after losing three-fifths of its purchasing power between 1914 and 1920, almost doubled in value between 1920 and 1932, when its value again began to fall. In comparison with nineteenth-century experience, this instability was abnormal. In Britain, for example, the maximum fluctuation in the value of money, one way or the other, was probably not more than 20 or 30 per cent. in the ninety years before 1914.

[1] It must not be supposed that the Board of Trade Index is a measure of the wholesale price of *British* produce. It is heavily weighted with imported foodstuffs and raw materials so that it is highly sensitive to changes in import prices. A fall in the index, therefore, may mean that British manufacturers are paying less for their raw materials, not that their products are fetching lower prices. Even the sub-index for "manufactured articles" does not provide a satisfactory measure of the wholesale price of British goods. For the price of manufactured articles which incorporate imported raw materials fluctuates with the price of these imported materials: the price of plain white cloth, which passes as a manufactured article, is closely dependent upon the price of imported raw cotton. The sub-index for manufactured articles fluctuates much more narrowly than either of the other sub-indices, the greatest fluctuations being, as we should expect, in the sub-index for basic materials.

Prices fell steadily, but very gradually, from the end of the Napoleonic Wars until 1896, with a short interval of rising prices in the fifties and sixties. After 1896 prices began to rise again, but until 1914 the rise was no more rapid than the previous fall. The course of prices in Britain since 1914 is illustrated in Chart I. Changes in the cost of living are shown in the Ministry of Labour's Index, and changes in wholesale prices, which were more pronounced than those in the cost of living, are shown in the Board of Trade's Index.

Why do these changes in the prices and in the value of money take place? Sometimes, when prices are rising rapidly, people blame "profiteers." But this is to confuse cause and effect. It is not because producers have become greedier that they charge higher prices; it is because prices are higher that producers are able to earn abnormally large profits. This is obvious enough when prices *fall;* low profits are the consequence, not the cause, of low prices. But when prices begin to rise our natural anxiety to find a scapegoat, and our detestation of excessive profits made at our expense, blind us to the deeper forces making for inflation, of which profiteering is only a symptom. A less simple-minded explanation is that goods become dearer because they are scarcer. In war-time, for example, imports of foodstuffs are cut off, and fewer workers are available for ordinary industrial employment; so prices rise. In peace-time, on the other hand, prices tend to fall because technical progress enables more goods to be produced. Even this explanation, however, is not very satisfactory. Prices are just as likely to rise because there is more money to spend as because there are fewer goods on which to spend it. Like the value of everything else, the value of money depends upon how much there is of it. This is a truth which we recognise every time we use the word "inflation." For most of us (including the Oxford Dictionary) are not quite sure whether we mean by inflation a rise in prices or an increase in the quantity of money; and we could hardly be in doubt unless we recognised that these two things were somehow connected with one another.

**The Quantity Theory of Money.**

The connection between prices and the quantity of money was first given clear expression in what has come to be known as the Quantity Theory of Money. Briefly, the theory is that the price level is governed by two flows—the flow of money on to the market for goods, and the flow of goods coming forward for sale against money. The larger the flow of money and the smaller the flow of goods, the higher will be the price level and the lower the value of money. Now the flow of money, which is simply our money outlay or expenditure, depends upon how many units of money we have, and the rapidity with which this money circulates. The flow of goods depends upon the volume of current production, and upon changes in stocks of finished goods. There are three possible reasons, therefore, why the price level may change between one year and another—first, because of a change in the quantity of money in our

possession; second, because of a change in the velocity of circulation of the stock of money; and third, because of a change in the volume of output or of stocks. The first change is on the side of supply; the second and third are on the side of demand. If money circulates less rapidly, this must be because we are turning over our stock of it more slowly and are using more of it *in relation to* the money transactions in which we engage; the demand for money rises. If the volume of output increases, more goods have to be exchanged; there is more work for money to do, and again the demand for money rises. This increase in demand has exactly the same effect as a reduction in supply in raising the value of money (or in bringing down prices). If demand remains constant—if money circulates with a constant velocity against a constant volume of goods—the price level will vary with, and in direct proportion to, the quantity of money.

The Quantity Theory is sometimes thrown into the form of an equation. If for the supply of money we write M; for the velocity of circulation, or average number of times a unit of money is spent, we write V; for the volume of transactions or of goods exchanged we write T; and for the price level we write P—then it is easy to show that:—

$$MV = PT.$$

For MV is our money outlay. If, at any given time, everyone has £20 in his possession (M) and if every £1 is spent ten times on the average (V) during a period of, say, one year, then the average man's money outlay in the year is £200 (MV). PT, being the quantity of goods sold, multiplied by their price, is the value of sales. But sales and purchases must be equal: the value of what is sold to us (PT) is equal to what we have to pay (MV).

The equation can be re-written so as to read:—

$$P = \frac{MV}{T} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and this formula, which we owe to Professor Irving Fisher, is generally referred to as the Fisher formula. A rather different formula—the Cambridge formula—is arrived at by considering, not the *flow* of money and of goods, but the *stock* of money which people wish to *hold* at any point in time. People carry about with them or hold at the bank a certain quantity of purchasing power—enough to buy, say, K units of goods. The value of K units of goods, or PK, is simply the demand for money; and since the demand for money must be equal to the supply, we have:—

$$PK = M, \text{ or } P = \frac{M}{K} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

If the quantity of money increases, and people want to hold the same amount of real purchasing power as before (i.e., if K is constant), the price level will be levered upwards. At the higher level of prices, K units of goods will cost more to buy, and people will, therefore,

require to hold more money. The increase in the supply of money will be matched by the increase in the demand which is brought about by the rise in prices.

**The Demand for Money.**

The Cambridge equation draws our attention to the fact that we want money for its own sake; that we want to hold a store of money just as we want to hold a store of clothes; and that we are put to just as much inconvenience when we find ourselves without money (in the sense of cash) as we should be if someone found us bathing and made off with our clothes. The sum of our individual stores of money makes up the total supply of it, and if more money is put into circulation, either by the government or by the banks, then *somebody* must increase his store. All the money in the country must, at any given instant in time, be in someone's possession. Whoever holds money, moreover, must presumably have some motive for doing so, since he could, if he chose, earn interest by lending it. What are the motives which prompt people to hold money? Upon what, in other words, does K depend?

There are, broadly speaking, three sets of motives:—

(1) **The Receipts and Payments Motive.**—First, we may hold money in order to bridge over the gap between the time when we receive money and the time when we pay it away. We receive our incomes in the form of money, as wages, salary payments, dividends, and so on, and pay them away in exchange for goods and services of all kinds. If receipts and payments were simultaneous—if, for example, we ran up accounts and paid them off as our income reached us—we should require very little money. But if we are paid only once a quarter, and have payments to make at a steady rate, we shall be left at the beginning of the quarter with a large cash balance which gradually diminishes until it is replenished at the next pay-day. If we are paid weekly, our income will reach us in smaller lumps and we are not likely to hold so large a balance. The time intervals which separate income payments, therefore, have an important influence on the amount of money which the community normally holds. The same is true of the time intervals which separate payments for goods and services. If income tax, for example, is paid half-yearly it will be necessary to build up a cash balance every six months so as to be able to pay the tax, whereas a tax payable weekly would not add very appreciably to the normal stock of money held by the taxpayer.

The gap between receipts and payments is not enough of itself to prompt consumers to hold money. The gap must also be so short that it is not worth the consumer's while to lend his spare cash and earn interest on it. If the interest which he could earn is small (e.g., on deposit account with a bank), the mere trouble of arranging a loan may be counted too great in comparison; and even if the interest is appreciable (e.g., on industrial debentures) the cost of making the loan (commissions on sale and purchase) may cancel

most of the prospective gain. It should also be kept in mind that the payments against which the consumer holds money include payments for investments which he proposes to make later. A man may resist the temptation to lend, even temporarily, the small sum which he has accumulated, because he hopes shortly to have enough for a larger and more profitable investment.

The same time intervals between receipts and payments which induce consumers to hold a stock of money exist in business. Business men, too, therefore, hold money balances to provide for any large payments which are in prospect or as the residue of receipts not yet expended. Indeed, in an industrial community, the balances held for business purposes may be just as large as the balances held by consumers against future payments. The cash requirements of business will vary with trade activity, rising in a boom and falling in a slump. In a boom, when higher profits are being made, larger balances will be held to provide for increased dividend disbursements, and to finance extensions to buildings and plant. Other payments and receipts will also be higher in value, partly because of expanding trade, partly because of rising prices, so that, if the time intervals between receipts and payments remain as before, businesses will find themselves holding, or requiring to hold, more money. In a slump, all these tendencies will be reversed, and businesses will hold reduced money balances. Consumers respond to fluctuations in trade in the same way. The higher their income, the more money they will hold, and the higher prices are, the more money they will require to hold in order to make their normal purchases. When trade is good, and incomes and prices are rising, consumers will increase their stock of money; and in a depression, with incomes and prices falling, they will diminish it. Thus the strength of the first motive to hold money depends mainly upon the level of incomes and prices, and upon the normal time intervals between receipts and payments.

(2) **The Speculative Motive.**—The second motive is an expectation that prices will fall. This expectation will lead people to hold money instead of using it to buy goods which they hope to obtain later at lower prices. If the expectation is sufficiently widespread it will *cause* prices to fall and so produce its own justification. But the expectation cannot be universal or prices will fall at once into line with the general market expectation; for consumers will not pay prices which they think too high, and producers, sharing the general opinion, will have no motive to hold out for existing prices. The fall in prices will remove the divergence between fact and expectation, and unless *some* consumers expect a further fall, the second motive to hold money will disappear. In other words, the second motive presupposes devisions of opinion on the future trend of prices. One group of peoples satisfied with existing prices, continues to buy. Others, anticipating a fall, reduce their purchases and add to their bank balances instead. But prices at any given moment are always low

enough to clear the market of all goods offered for sale, an addition to stock being regarded as a purchase by the holder.

When prices begin to rise, and are commonly expected to rise further, the second motive to hold money is transformed into a motive to get rid of it. No one wants to hold money, a depreciating asset, longer than is necessary, and it is passed on hastily like a bad penny. In a major inflation, people keep by them just enough money to meet day-to-day expenses; and do all they can to reduce the normal intervals between receipts and payments. Thus the store of real purchasing power which they hold in the form of money (K) falls heavily. At the end of October, 1923, the note issue in Germany had increased 400 millionfold in comparison with 1913; but in spite of this enormous inflation, which raised the printing of notes to the status of a major industry, the real purchasing power of the note issue in terms of gold was less than a fortieth of the purchasing power of the pre-war note issue.

In normal times, when consumers have little reason to expect either a rise or a fall in the prices of *goods*, they may hold a considerable stock of money in the expectation that Stock Exchange securities will fall in price. They have to choose not only between holding money and spending it, but also between holding money and investing it. But they will not purchase an investment so long as they have hopes of a better bargain in the near future; they will prefer to keep their wealth liquid in the form of money until security prices come into line with their expectations. Here again these expectations cannot be universal, or prices would fall at once until there was no expectation of a further fall. There must be a division of opinion, prices reflecting the general state of expectation, while individual investors sell or buy according as they take a less favourable or more favourable view than the market as a whole. Such a division of opinion was particularly conspicuous on Wall Street in the summer of 1929 when very large purchases of securities were made by bullish speculators and correspondingly large sales by bears. The bears, having sold out, added to their bank balances, preferring to hold money until after the fall in security prices which they anticipated.

The strength of the second motive depends upon the firmness with which conflicting opinions are held about the probable course of prices, and particularly about the probable course of security prices. If prices are falling, and if there is a strong body of opinion which expects the fall to continue, there will be a large demand for money from those who refrain from buying goods. The expectation of falling commodity prices will communicate itself to the Stock Exchange, making money a more attractive investment (since its value is expected to increase) and securities a less attractive investment (since profits are expected to follow prices downwards).[1] Unless these unfavourable

[1] Bonds, on the other hand, will become a *more* attractive investment and will tend to rise in price. That is, the rate of interest tends to fall, other things remaining the same.

expectations are immediately discounted by a heavy fall in the prices of securities, bearish opinion will gain strength, and investors will increase their demand for money. If this happens, then the weakening in the demand for money from the first motive during a trade depression may be more than compensated by a strengthening in the demand from the second motive. But experience suggests that this is rather unlikely, and that a division of speculative opinion is greater at the *turning points* of boom and slump than during the periods when prices are steadily rising or falling. There is a hesitancy when prices first begin to rise, and a faltering of business sentiment when prices first begin to show signs of falling, both of which increase the demand for money. At other times the expectation of falling prices does not appear to lead to large variations in the demand for money.

(3) **The Precautionary Motive.**—The third and final motive to hold money springs from uncertainty. We can never be quite certain in what payments the future will involve us, nor what receipts we shall ultimately obtain. Lacking certainty, we arm ourselves with money against emergencies—against calamities like illness and death, and the expense to which they put us; against sudden calls on our charity; against opportunities of purchase on favourable terms; against delays guessed at or unforeseen in the receipt of income. The money that we hold gives us security; it makes us *liquid*. The motives to liquidity are particularly strong when it is the price of securities about which we are uncertain. Even if we do not expect that security prices will fall—that is, even when the second motive to hold money is absent—we may have too little confidence in our judgment to come in and buy. We may prefer to hold a large part of our wealth in monetary form until the future course of prices is less in doubt.[1]

The third motive to hold money is necessarily reinforced by the second. Differences of opinion about the future course of prices pre-suppose a state of uncertainty, and the greater this uncertainty, the more divided is opinion likely to be. Just as the second motive is strongest when the trend of prices is in course of reversal, so also is the third motive. But the third motive is also of importance at other times, since it does not arise solely from uncertainty about prices.

**Hoarding.**

A demand for money from the second or third motive is primarily a demand for money to hoard. We want the money, not for some specific purpose but as a safeguard or as the best investment open to us. For the time being we want to keep it idle. To suggest therefore—as is so often done—that there is something heinous about "piling up idle deposits in the banks" is to misunderstand completely the functions of money. It is the business of bank deposits—other than those held from the first motive—to be idle.

[1] Compare the discussion of this third motive above, pages 292-3.

Nor does it lie within the power of the general public to alter the total volume of bank deposits, except by devious ways such as stuffing their pockets with notes; the level of bank deposits depends upon the monetary policy of the banks themselves.[1] Nor, in point of fact, are bank deposits at an abnormally high level in the years of depression when "idle deposits" are most lamented; they are generally *below* the level reached in more prosperous years. What *is* true, and what *does* lend colour to the complaints, is that a larger proportion of bank deposits is hoarded in a depression—that is, is held exclusively from the second and third motives. But the remedy lies not so much with the public as with the banking system, which, by creating more money, can satisfy the increased desire for liquidity. The public might, of course, spend more, and give money more work to do, but this would change, not their aggregate bank balances but their savings.

A change in the demand for money, whatever the motive, can be measured in two quite different ways.

(1) In terms of the Cambridge equation we may ask: "Has there been a change in K? Are people holding a larger store of real purchasing power?" When we ask this question we are really asking whether prices have fallen. For if the *supply* of money remains constant—and no anxiety on the part of the public to obtain more money can of itself increase the supply—an increase in K is possible only if there is a corresponding fall in P. This is immediately obvious from the equation $M = PK$. Thus we are able to measure the change in the demand for money by measuring the change in the price level.[2]

(2) We may measure the demand for money in terms of the strength of our preference for it in comparison with alternative investments. We sacrifice interest by holding money, and this interest is, so to speak, the price which we pay for money. The greater our demand for money, the more interest we shall be prepared to sacrifice and the higher, therefore, will interest rates go before we are willing to lend money or buy bonds. This line of thought, if pursued far enough, leads to the conception of liquidity-preference.[3] The demand for money is seen to govern, not the price of commodities, but the price of bonds.

It is clear that the demand for money as measured in the first way (the Quantity Theory way) is not the same as the demand for money measured in the second way (the liquidity-preference way). There

[1] For a demonstration of this point, see page 326 *et seq.*

[2] Notice that the Quantity Theory regards the change in the demand for money as the *cause* of the change in the price-level, not just as equal to it. The above discussion, however, suggests that we might almost equally well explain the change in the demand for money as a *consequence* of the change in the price-level. This criticism will be developed later (page 353).

[3] See above, page 258.

are, however, links between the two. The first method of measurement leaves investment out of account and concentrates on the receipts and payments motive for holding money; the second motive leaves spending out of account and concentrates on the speculative and precautionary motives for holding money. Moreover, an increase in liquidity-preference, and hence in interest rates, can be shown to act as a brake on business activity, bringing down prices, and increasing K; both kinds of "demand for money" tend to move together. The demand for money in the first sense is a valuable conception in broad comparisons of monetary habits over long periods of years or at varying price levels; the demand for money in the second sense is a conception indispensable to an analysis of fluctuations in prices.

The terms "K" in the Cambridge equation and $\frac{T}{V}$ in the Fisher equation are different expressions for the same thing. Suppose, for example—what is roughly true of Great Britain—that consumers, on the average, try to hold a month's income in the form of money. Then, if we measure T and V over a period of a year, T is equal to the real National Output and V is equal to 12 (since in spending our income we must make use of each pound 12 times on the average). But K, the command over goods which we hold in the form of money, is one-twelfth of the real National Output; in other words, it is equal to $\frac{T}{V}$. It follows that so long as T remains constant, the same forces govern K and V; the demand for money is inversely proportional to the velocity with which money circulates. It follows also that the formulæ $P=\frac{M}{K}$ and $P=\frac{MV}{T}$ are interchangeable with one another.

The two formulæ have the appearance of being very precise and definite. But when we come to interpret them, difficulties arise. First of all, the symbols are ambiguous. Take P, for example. There is no one price level to which P *must* refer; there are, as we have seen, many sectional price levels, depending upon the particular group of commodities which money is used to buy. We might try to make P refer to the price level of *all* commodities and transactions. But this "general price level" would be a hotch-potch of no real interest. We cannot average, in any way that makes sense, movements in wholesale and retail prices, and in the prices of raw materials, Stock Exchange securities, old houses and motor-cars, skilled and unskilled labour, and all the other goods and services on which money is expended. What we can do—and what we should do—is to make P refer unambiguously to the price level of finished goods—the goods on which, as consumers, we spend our income. T, M and V can then be defined correspondingly. T will include sales of finished goods only; it will not include sales of raw materials, semi-finished goods, Stock Exchange securities, etc. M will include money held by income-receivers with a view to disbursement on goods and

services; it will not include money held for business purposes or hoarded as an investment. V will be the velocity of circulation, not of all units of money, but of those units only which we include in M. On this definition, V is equal to the number of times that a unit of money held by income-receivers circulates from hand to hand in the course of a year, and it is referred to, therefore, as the *income*-velocity of circulation.

The dependence of the price level on the quantity of money is clearly less direct than the simplicity of the Fisher formula might seem to suggest. It is not the *total* amount of money (in coin, notes and bank deposits), nor the activity of this total that exerts pressure on the price level, but the amount and activity of a portion only of the money supply. How great this portion is and what changes take place in its size from time to time it is not easy, in the absence of reliable data, to say. We certainly cannot assume that every increase in the total money supply increases M proportionately. It may well be that a large increase in the money supply adds little or nothing to the amount of money in circulation against finished goods and is simply hoarded by consumers in preference to alternative investments. In times of trade depression, indeed, this reaction to an increase in the money supply is typical rather than exceptional; the increase tends to be absorbed by investors and by business men who are anxious to improve their liquidity. Thus the change in the price level is by no means in proportion to the change in the total quantity of money.

A second difficulty in interpreting the Fisher formula is that while it tells us that P is *equal* to $\frac{MV}{T}$, it does not tell us what causal connection there is between M, V, T and P. If we were to re-write the formula to read $M = \frac{PT}{V}$ we might be tempted to deduce that P, T and V summed up the forces governing M. But we should be hopelessly wrong. For we know that M is determined by the banking system and by the government—that is, by forces which do not appear in the formula at all. Similarly V, although it is *equal* to $\frac{PT}{M}$, is determined primarily by the monetary habits of consumers; and T, although it is equal to $\frac{MV}{P}$, depends mainly upon producers' expectations of profit. An algebraic formula, therefore, is a very different thing from a statement of causal connection. The Fisher formula gives us no more reason for supposing that P depends upon M, or V, or T, than for supposing (what is by no means impossible) that M and V and T depend upon P.

Nevertheless the Quantity Theory does postulate a causal connection between the variables in the formula. The whole point of the theory is that an increase in M *causes* an increase in P—that an increase in the money supply tends to spend itself in driving up the cost of living

rather than in making us better off by causing more goods to be produced. If there is no such connection of cause and effect—if it is the level of output and not of prices which rises when more money is created—then the Quantity Theory is misleading and mischievous. For nothing is easier than to increase the quantity of money, and it would be criminal folly to allow a mistaken theory to deter us from so inviting a remedy for human poverty. Hence it is important to consider, in the light of experience, what effects—on the price level, on trade, and on the velocity of circulation—an increase in the money supply is likely to have. When we go to the facts, what causal connections between M, V, P and T do we observe?

Broadly speaking, the generalisation made in the Quantity Theory is vindicated by experience—a persistent increase in the quantity of money does tend to raise prices and lower the value of money. Not only so, but there is no other force which, in the course of human history, has made so strongly for a fall in the value of money. Any sudden fall, such as takes place in war-time, or any prolonged fall, such as took place in the sixteenth and seventeenth centuries, is almost invariably to be traced to inflation of the money supply. In looking for the main causes of rising and falling prices, therefore, we are right in singling out changes in the quantity of money for special emphasis; and we are right in stressing that an increase in the money supply, if too large or too long-continued, brings no lasting benefit in increased output and simply raises the price of what is already being produced. But we must not push these conclusions too far. There are times when quite large changes take place in the price level without any accompanying change in the quantity of money; and there are times when an increase in the quantity of money helps to stabilise or to expand the output of industry. In the swing from industrial depression to industrial prosperity and back again, prices generally rise and fall. The money supply *may* undergo similar changes, the banks lending more money in the upswing and contracting credit in the slump. But even if the money supply were to remain constant throughout, booms and slumps would not disappear. Here, therefore, changes in the money supply are only one cause of price fluctuations amongst several; and, as we shall see in a later chapter, the Quantity Theory is of little assistance to us in trying to understand such *cyclical* fluctuations. In times of industrial depression, moreover, it ceases to be true that an increase in the money supply is of no real advantage to society. The first effect of creating and spending more money in a slump is to generate a demand for goods which were previously unsaleable and to bring back into employment workers who were previously unemployed. These workers in turn, out of the higher incomes which they earn in wages, can buy more goods and create opportunities of employment for other workers. So long as there are still large numbers of men out of work, and large stocks of raw materials available, there is no reason to expect a rise in prices. The burden of adjustment to the increase in M falls on T, not on P. There is an increase in the

supply of goods to match the increase in the supply of money and with these additional goods a higher standard of living can be maintained.

Suppose, however, that money continues to be created and spent until a shortage of raw materials or of labour develops. We then enter upon a period of real inflation to which the Quantity Theory can be applied in its full rigour. A mounting stream of money expenditure, fed by additions to the money supply, meets a constant stream of goods. Prices begin to rise, and the rise is soon accelerated by an increase in the velocity of circulation. Finding that money is losing its purchasing power, consumers hasten to exchange it for goods; they buy clothes, furniture, jewels—anything that can be relied upon to keep its value. These purchases force up prices still further, and the rise in prices, in a kind of vicious spiral, causes money to circulate faster than ever. The rise in prices reacts also on the supply of money. A government which resorts to inflation as the easiest method of balancing its budget soon finds that while most of its expenses move upwards with the price-level (interest payments apart) tax receipts lag behind. The more notes it prints one year, the more notes it needs the next. Once well started on the path of inflation governments have every inducement to continue along it. The same is true when inflation is initiated, not by the government, but by the banking system. By granting credit too freely the banks can add just as effectively to the money supply as the government when it prints too many notes; and just as rising prices put pressure on the government to issue still more notes, so also pressure is put on the banks to grant still more credit. Rising prices mean easy profits; one need only borrow money and use it either to buy or to manufacture stocks of goods, selling them at the higher prices which they can be relied upon to fetch so long as the upward movement continues. The banks, therefore, are besieged with borrowers; some who are in search of a speculative profit, and some who, at the higher level of prices, cannot obtain enough working capital for their business unless they are given a proportionately higher loan. The more credit is created, the more rapidly do prices rise, and the more urgent is the clamour for still more credit. It is a familiar paradox that money is never so scarce as when it is most abundant.[1]

The initial reaction of inflation on the volume of output is favourable. Trade becomes brisker under the stimulus of rising prices and profits, and unemployment falls away to a low level. So long as the supply of money keeps on increasing and the rise in prices is moderate, production continues at a high level, at or near capacity. But if the inflation ends, or if it gets out of hand, industry is disorganised and

[1] It is amusing to read of the apology made by the Reichsbank, at the height of the German inflation, for its inability to meet the demand for notes, and of its hope that, by increasing output to a trillion marks per week, it would catch up with the demand. (C. Bresciani-Turroni: "Economics of Inflation.")

output begins to fall. A check to inflation forces producers to revise their views of the probable course of prices and wipes out the profits of those who have discounted the future too confidently. Producers can neither hope to make as large a profit as before nor be so sure of making any profit at all, and they will have a double incentive, therefore, to curtail output. If inflation proceeds too rapidly, the repercussions on output may be equally unfortunate. It is so easy to make a profit that the competitive weeding-out of inefficient firms and methods is almost completely suspended and productivity falls heavily. There is a multiplication of middlemen and a great waste of ingenuity. The trend of prices becomes increasingly incalculable, planning ahead for any considerable period becomes impossible, and at the height of the inflation, unemployment begins to increase.

$P=\frac{MV}{T}$ is not the only formula in terms of which we can represent the forces governing the price level. From some points of view it is not even the best formula. Its advantages are that it concentrates our attention on the force which, historically, has dominated the course of prices; that it hangs out a warning against the fatal expedient of inflation by reminding us that money loses its value whenever we make it too plentiful; and that when we are making broad comparisons between price levels in different years, it is a useful pointer to the forces underlying the change from one price level to the other. Its weakness is that it does not provide us with an adequate picture of the *process* of change. We can see in a general way that changes in the quantity of money will work through to the price level; but if we want to make a closer study of the machinery of change we find ourselves beginning to discuss wage rates, interest rates, expectations of profit, and other variables which find no place in the formula. We can still construe everything—with some difficulty—in terms of the formula; but only when we have worked out, by another route, what is happening. We do not learn from the formula; we read into it what we have already discovered.

The formula is, in fact, too simple. It seems a natural application of the theory of value—the greater the supply of money, the less its value. But what from one point of view is a change in the value of money, is from another point of view a change in the price level of commodities. Now a change in the price level of commodities can come about only through a mass of changes in the prices of particular commodities. And if we ask of the theory of value why the price of any *particular* commodity should change we receive an answer which has apparently no connection whatever with the explanation given by the Quantity Theory why prices in general should change. We learn, for example, that the price of a commodity will rise if there is an increase in its marginal cost of production, or in the demand for it, or in the monopoly power of producers. This opens up quite a different line of thought from the Quantity Theory. It suggests four reasons why the prices of commodities may rise and the value of money fall.

First, there may be a general rise in efficiency wages; this will raise costs and force producers to charge higher prices. A rise in wages commonly takes place in the concluding stages of an industrial boom, when the bargaining power of the trade unions is strengthened by a general shortage of labour. Employers are less disposed to contest demands for wage advances because they are then making abnormally large profits; and public sympathy is generally on the side of the workers because the rising cost of living is eating into the real purchasing power of existing rates of wages. The rise in wages, coming at a time when prices are already moving upwards, accelerates the upward movement. Each fresh rise in prices will provoke renewed demands for wage increases, and these, if granted, will raise prices still further. It might seem as if successive increases in wages and prices could follow one another without limit—and in an inflationary spiral the two do feed on one another like Kilkenny cats. But a persistent rise in prices or in wages is possible only if the money supply keeps in step. A rise in wages increases the demand for money; the worker, having earned more money in wages, tends to keep a larger stock of it by him or at the bank; and the employer, having more money to pay in wages, requires to hold a larger balance to cover his wage bill. The consequent rise in prices works in the same direction; since goods are dearer, people need more ready money in order to make their usual purchases, and business men, if they wish to lay in stocks or to engage in dealings of any kind, find it equally necessary to keep a larger bank balance. If the money supply were to be kept constant in face of this increased demand for money, interest rates would be forced up, there would be a set-back in trade, and prices would come tumbling down, dragging wages after them. Other things remaining the same, an increase in the money supply is indispensable to an increase in wages and prices. Here, therefore, we have a link with the Quantity Theory—an increase in the quantity of money may not *cause* a rise in prices; certainly it has no obvious and direct influence on efficiency wage rates, but it is a *pre-requisite*.

Secondly, there may be an expansion in output under conditions of increasing marginal cost; the increase in marginal cost will naturally communicate itself to price. Here we have the explanation of the gradual rise in prices which normally accompanies a trade revival. Workers of less than average efficiency are taken on, old and obsolete machinery is brought into use, organisation becomes slacker, and employers resort to a variety of costly makeshift devices for increasing output, such as working overtime and running machinery beyond capacity. The result is a rise in marginal cost—although initially, at low levels of output, there may be a slight fall. The increase in cost is passed on to the consumer, and prices rise as recovery proceeds.

Thirdly, there might be an increase in the monopoly power of producers either in hiring the factors of production or in disposing of their product; this would allow them to charge a larger margin

of profit into the price of their goods and force up prices relatively to costs. It is doubtful whether much importance can be attached to changes in monopoly power, but it seems likely that, so far as they go, they aggravate price fluctuations. Competition between producers is keener when prices are falling than when they are rising; in a slump, therefore, increased competition causes prices to fall more than in proportion to marginal costs, while in a boom the relaxation of competitive pressure allows producers to raise prices more than in proportion to marginal costs.

Finally, a rise in prices might be due to an increase in the demand for goods and services—that is, to an increase in expenditure. Such an increase in expenditure, as we have already seen in discussing the Quantity Theory, will raise prices because no commensurate increase in the supply of goods is likely to result from it. But how, in the absence of a change in incomes, can an increase in expenditure take place? How are people able to spend more if they have no more to spend? The explanation offered by the Quantity Theory is that more money is created or the existing stock circulates more rapidly. Expenditure, that is, is broken up into M and V and the influence of each on the price level is studied separately. In modern economic theory, expenditure is divided up in a quite different way into outlay on consumable goods and outlay on capital goods. The first outlay is comparatively stable, the second highly variable. Our expenditure on consumable goods is a thing of custom and habit, ruled by familiar standards, and altered only when our income alters or when the incentives to thrift gain or lose in strength. Our expenditure, either as producers or as consumers, on capital goods—on factory buildings, ships, machinery, dwelling-houses, motor-cars, etc.—proceeds by fits and starts, for it is not tied, like purchases of food and clothing, to present needs of a constant urgency, but is sensitive to our changing expectations about the future. Capital outlay is outlay to meet the wants of the future; a dwelling-house, for example, may last for hundreds of years. The longer the period for which capital has to be sunk, the more difficult will it be to estimate whether the capital outlay will yield a profit, and the more responsive will expenditure be to changes in the business outlook. We need not, at this stage, enumerate the factors which go to determine capital outlay nor all the consequences of its instability; it is enough to appreciate that we have here something explosive which can send prices rocketing. For it is changes in capital outlay that are the real driving force behind changes in the price level. Changes in wages, output and monopoly power play the part of first-line reserves in general price movements; they follow and reinforce, but rarely take the lead.

The sequence of changes resulting from an increase in capital outlay we shall study in more detail later. Briefly the effect on prices is this. The increase in outlay, say on housebuilding, tends first of all to raise the price of houses because of the increased demand for building materials and labour, and because of the higher profit margin which

builders are able to charge. Next, more is spent on consumable goods by workers who were previously unemployed and by investors who find themselves making larger profits. This increased demand is supplied from stock and renewal orders are sent to manufacturers—the more willingly since trade appears to be reviving. Manufacturers increase their output to meet these orders and, finding marginal costs beginning to rise, put up the price of their goods. Meanwhile demand is continuing to increase because the expansion in output has put more money in the pockets of the additional workers employed. Once recovery is well under way, claims will be put forward for advances in wages, and if these are granted the rise in prices will be given fresh impetus. At the same time, the increasing anxiety of consumers to obtain supplies will add to the bargaining power of sellers and raise prices still further. The rise in prices and the improvement in the business outlook will encourage producers to increase their capital outlay (for example, to finance extensions of plant), and the same series of repercussions will be felt all over again.

It remains to add that changes in capital outlay may result from changes in interest rates and credit conditions, and that both of these are likely to accompany changes in the quantity of money. A fall in interest rates will give producers an incentive to borrow more, and their expenditure out of borrowing will increase capital outlay. The fall in interest rates may itself be due to an increase in the supply of money.[1] The money supply does exert pressure on the price level, therefore; there is no contradiction between the Quantity Theory and the modern approach to monetary problems. But the connection between money and prices through interest rates is more roundabout than the Quantity Theory seems to imply. There are links in the chain of connection which are hidden from us in the formula $P = \frac{MV}{T}$.

### The Consequences of Changes in the Value of Money

**(1) On Distribution.**—Incomes, as we saw in Chapter 21, are either fixed by contract or variable with market prices.[2] The income of labour, supposing no change in employment, is fixed by a number of wage bargains, the majority of which can only be revised slowly and after much bickering. Income from loan capital is fixed for periods varying with the duration of the loan. Income from the ownership of land and of dwelling-houses is also slow to change, either because rents are fixed for a number of years ahead or because custom inhibits the raising of rents to existing tenants. Profit, on the other hand, is free to vary as prices move up or down. A change in the value of money, therefore, effects a transfer of income from one income group to the other. If prices rise, wages lag behind until

[1] *cf.* above, pages 316-7.

[2] See page 272.

workers can bring pressure to bear on their employers for an increase —generally too late to maintain real wage rates. The rentier receives in interest an income which is constant in terms of money and steadily falling in terms of the goods which it will buy. The landlord finds his real income similarly reduced. The employer, on the other hand, benefits from the rise in prices in increased revenue, and as there is no comparable increase in outgoings, pockets a windfall profit. Nor is it only income which is transferred to him. For since debts can now be discharged in money which has fallen in value, the real burden of debt is diminished, so that there is a transfer of wealth from creditors to debtors. The employer who is working with borrowed capital is thus the chief beneficiary of a fall in the value of money; while the rentier, who is most helpless to protect himself by seeking a revised loan contract, is the chief victim. When prices fall, all this is reversed. The real burden of fixed charges is increased, adding automatically to the real income and wealth of the rentier, and crippling the employer who is in his debt. Wages and rents again lag behind prices, the fall in prices conferring a bounty on workers who keep their jobs and on landlords who keep their tenants.

(2) **Production.**—Production is governed by producers' expectations of profit. Rising prices, therefore, by adding to profits, stimulate production. But since production cannot expand beyond the limit set by full employment of labour, the stimulus is felt only in the early stages of an inflation when productive resources are still idle. Thereafter, moderate inflation is likely to maintain output at a higher level than steady or gradually falling prices. This effect, however, is not altogether certain. Easy profits relax the pressure on employers to maintain efficiency at its highest pitch; it is not when trade is most active that the greatest advances in technique and organisation are made. Thus we must set a fall in productivity against the rise in employment. At what stage the one will counterbalance the other it is impossible to say. But it is clear that if prices rise too rapidly, there will be no further gain in employment (there may even be a fall) and a considerable drop in productivity.

Falling prices, unless accompanied by falling costs of production, mean falling profits and unemployment. Producers are unable to cover their costs and try to protect themselves by curtailing output. Part of the burden is thus passed on to the wage-earner, who loses his job or is put on short time. A smaller quantity of goods is produced so that the National Output (and hence the National Income) is diminished. Deflation, therefore, is generally more creative of social distress than all but the most rapid inflation. If inflation tends to wipe out savings and inflict hardship on special classes, deflation can be far more disastrous in causing unemployment amongst those whose resistance to it is the weakest, in driving businesses into unmerited bankruptcy, and in impoverishing the community by throwing production resources into needless idleness.

Inflation and deflation have each their peculiar evils. Some economists prefer the one, some the other, and some a little of both. Where there are such differences of opinion, all soundly defended and suitably qualified, the layman is to be pardoned if he calls for "A plague on both your houses!" and decides that a steady price level is best of all. We shall see later, however, that the dilemma is an unreal one, and that monetary policy should be not so much directed to making prices behave in one way rather than in another, as towards more fundamental objectives such as the cure of unemployment. Rising or falling prices may be the necessary consequence of a given policy, but are not its immediate objective.

## CHAPTER 27

## MONEY AND FOREIGN EXCHANGE

### The Balance of Payments

PAYMENTS between countries generally differ from domestic payments in that they involve the exchange of one kind of money for another. Each country has its own money, which is not freely accepted abroad, but must be changed into foreign money at a rate which may be either fixed or variable. The rate of exchange, whether fixed or variable, is a price; and, like any other price, must be such as to balance supply and demand. At any given rate, sales of foreign exchange must be equal to purchases; the payments made by foreign countries must be equal to the payments made to them. If the balance of payments is tipped against a country, there will be increased pressure to buy foreign exchange, and, if the rate of exchange is free to vary, the price of foreign moneys will rise—that is, the currency of the country will depreciate. If the balance of payments is favourable to the country, there will be increased pressure to sell foreign exchange, and this pressure will tend to raise the value of the country's currency; or, in other words, cause it to appreciate.

This idea of a balance of payments which may be either favourable or unfavourable is fundamental to the whole theory of foreign exchange and requires careful explanation. The balance of payments may refer to any one of three things:—

(1) **The Balance of Trade.**—In speaking of an unfavourable balance of payments we may have in mind an excess of imports of merchandise over exports. Since details of these imports and exports are published monthly in the Trade Returns, and since merchandise forms a large proportion of total trade, this "balance of trade" figures very prominently in most popular discussions. But that it is of limited importance is clear from the fact that Britain has had an unfavourable balance of merchandise trade for over a hundred years without any unfortunate consequences. The balance is, in fact, only a partial balance since it omits some large "invisible" items.

**(2) The Balance of Payments on Current (or Income) Account.**—In addition to merchandise exports and imports there are various "invisible" exports and imports whose value is not accurately known and which are not included in the Trade Returns. For example, large sums are earned every year by the British mercantile marine in carrying goods and passengers from one foreign country to another, or between British and foreign ports. The Board of Trade records the value of British exports "f.o.b." (free on board—that is, at the port of embarkation), and the value of British imports "c.i.f." (cost, insurance, freight—that is, at the port of landing). Thus its returns exaggerate the adverse balance by failing to take account of shipping earnings. Since foreigners require to make additional payments to us to cover sea transport charges on exports carried in British vessels, and since we are spared from making payments to foreigners to cover sea transport charges on imports carried on British vessels, the earnings of British ships in carrying goods to and from Britain go to reduce any visible adverse balance of trade. Further deductions must be made in respect of the fares of foreign passengers in British ships, transport of goods between foreign ports, and disbursements of foreign ships and sailors in British ports. On the other hand, disbursements by British ships in foreign ports (for example, for fuel or in payment of port dues) add to any unfavourable balance. Shipping freights are a very large item in British trade—normally about £100 millions. Shipping, in fact, has for many years been our leading export industry.

A second "invisible" item is interest on foreign investments. A large part of Britain's imports—say, half of our total imports of "food, drink, and tobacco"—come in payment for part investments made by Britain in foreign countries. Interest payments to us are a kind of "invisible" export; they can be regarded as payments for the hire of British capital just as freight payments are payments for the hire of British ships. Similarly, interest payments on capital invested in Britain are "invisible" imports.

To these two items may be added a large number of others. We have payments to make to foreign countries on account of foreign films exhibited in this country, the disbursements of British tourists abroad, remittances to friends and relatives in other countries, the cost of upkeep of embassies and consulates, and so on. Against these payments must be set similar payments due to Britain by foreign countries, including the commissions earned on our marine underwriting, insurance and acceptance business, the pensions of Indian Civil Servants, etc. Once all these items have been included we reach a quite different balance, representing the difference between the value of goods and services currently produced in this country and sold abroad, and the value of foreign goods and services bought from abroad. This balance, which may be called "the balance of payments on income account," was consistently favourable to Britain all through the second half of the nineteenth century and did not become unfavourable (apart from the war years 1914–18) until 1931. Thereafter the balance on income account was not markedly favour-

able or unfavourable, but oscillated between small positive and small negative quantities until the outbreak of war in 1939.

(3) **The Balance of Short-Term Indebtedness.**—The balance of payments on income account has no necessary significance; it merely indicates whether we are lending to foreign countries or borrowing from them. If we sell more goods and services than we buy, then other countries must "pay" the difference by borrowing from us; while if we buy more than we sell, then we must borrow the difference from foreigners. Now there is nothing necessarily undesirable or "unfavourable" about raising capital abroad. Suppose, for example, that a country is engaged in building a new railway system and that its own savings are inadequate to finance the work. By borrowing from other countries it can make purchases of railway equipment abroad and pay for these purchases not by equal exports, but out of the borrowed funds. In the same way, if Britain has an unfavourable balance on income account, this may simply reflect unusually active investment at home; we may be drawing on our accumulated foreign investments in order to finance a building boom or the construction of armaments factories. A country can choose between investing its savings at home or abroad. The more it inclines to home investment the less favourable will be its balance of payments as measured by the difference between exports and imports; for it will have less resources to spare for the manufacture of exports and it will be more anxious to draw on the resources of other countries by importing from them.

Although there is nothing necessarily unfavourable about a negative (or "passive") balance of payments which is covered by *long-term* borrowing abroad, a negative balance which is met either by short-term, "makeshift," borrowing, or by exports of gold, is quite another matter. Long-term borrowing designed to finance the construction of capital assets can go on with impunity for very long periods. But short-term foreign loans are generally obtained only if there are good prospects that the loans can be repaid. A country can borrow on short-term to tide over a temporary deficit in its balance of payments; but it cannot continue to add to its short-term borrowings without simultaneously taking measures to correct the unfavourable balance. Similarly, it can finance a deficit in its balance of payments by the sale of gold; but, since its stock of gold is limited, the loss of gold cannot go on indefinitely. Short-term borrowing and gold exports, therefore, are palliatives which, so long as they continue, keep the rate of exchange from depreciating, just as the sale of goods from stock keeps their price from rising when demand increases. Both expedients are danger signals as well as palliatives. They indicate, that unless action is taken to correct the unfavourable balance and so reverse the inward flow of credits or the outward flow of gold, depreciation is inevitable.

The balance of payments in this sense may be unfavourable even when the balance on current account is favourable. A country may overlend; that is, it may lend or invest abroad more than its favourable

balance on current account and find itself losing gold or reborrowing on short-term. The experience of Britain between 1925 and 1931 showed that this is just as dangerous a situation as one in which a country borrows abroad both on long-term and on short-term. It is the balance of short-term indebtedness (including transfers of gold and of liquid funds) that governs the pressure on exchange rates.

The consequences of an unfavourable balance of payments differ according as a country is on the Gold Standard or not. When countries are on the gold standard, their exchange rates can vary only within narrow limits; the Central Bank has to offset the pressure of the exchange market by appropriate purchase or sales of gold according as the balance of payments is favourable or unfavourable. When countries are not on the Gold Standard, the limits to fluctuations in exchange rates are removed and the balance of payments acts on exchange rates without check. Even when a country is not on the Gold Standard, however, arrangements may be made to "peg" or to steady exchange rates by setting up Exchange Equalisation Funds. These Funds, which normally hold large amounts of gold or of foreign exchange, can intervene (like central banks under Gold Standard conditions) to prevent depreciation by supplementing the supply of foreign exchange and to prevent appreciation by making purchases of foreign exchange. We shall return to the operations of Exchange Equalisation Funds once the workings of the Gold Standard have been explained.

**The Gold Standard.**

A country is on the Gold Standard when its currency is convertible into gold at a fixed price. If gold coins are in circulation and all other kinds of money, such as bank notes, are promises to pay gold, then this condition is automatically fulfilled. Even if, as in Britain between 1925 and 1931, there is no gold currency but only a law instructing the Central Bank to purchase all gold offered to it at a fixed price, and to sell gold bullion at a slightly higher price, the condition is still fulfilled.[1] It is also fulfilled if the Central Bank is instructed to buy and sell at fixed prices not gold, but the money of some other gold standard country; for this money is itself convertible into gold at a fixed price. The first kind of standard was the one generally adopted up till 1914, the second, the Gold Bullion Standard, came into use after the war; while the third, the Gold Exchange Standard, was adopted by countries like India which could not afford a full gold standard, preferring instead to link their currencies with sterling, and so, indirectly with gold.

The Gold Standard is a means of preserving constant, or almost constant, exchange rates between different countries. If the United

[1] Under the Gold Standard Act of 1925 the Bank's buying price was fixed at 77s. 9d. per oz. of standard gold, and its selling price at 77s. 10½d. per oz. of standard gold in the form of bars of 400 oz. of fine gold. Since standard gold is 11/12 fine, this selling price was equivalent to 85s. 0d. per fine ounce in quantities worth not less than £1,700.

States and Great Britain are both on the Gold Standard, the dollar and sterling prices of gold will both be fixed, and the terms on which pounds and dollars exchange will, therefore, also be fixed. Suppose, for example, that the mint price of gold in New York is $35 an ounce and in London £7 an ounce. Then the mint par of exchange will be $5 to the £1. It will be foolish for an American to offer more than $5 for £1 since by buying gold and shipping it to London he can obtain pounds at this rate. On the other hand, no American will sell pounds for less than $5, since by buying gold in London and bringing it to New York he can convert his sterling balances into dollars at this rate. But the shipping of gold in either direction puts him to some expense, and various obstacles may be put in his way by central banks which do not like to lose gold too easily. If he can buy pounds on the market for foreign exchange, therefore, he will be willing to pay rather more than $5 to save himself trouble and expense, and if he has pounds to sell he will be willing to accept rather less than $5 for the same reason. Thus the Gold Standard does not entirely eliminate variations in exchange rates; it merely sets narrow limits to them. For each country there will be a maximum price of foreign exchange above which gold begins to flow out ("gold export point") and a minimum price below which gold begins to flow in ("gold import point"). The distance between these points—"the gold points"—depends upon:

(i) The cost of transport of gold from one financial centre to another, and the time taken to make the transfer. This item includes freight and insurance charges and loss of interest in transit.

(ii) The margin between the buying and selling prices for gold in the two centres. If the central banks in the two countries offer a lower price for gold than they charge for it, it will require a larger movement in exchange rates to make gold flow either in or out.[1]

(iii) The charges of the Central Banks for seigniorage, minting, etc., when gold coins are in use, the unofficial pressure which they can bring against the export of gold, and their ability to vary their buying or selling price by expedients such as altering the fineness of the gold bullion in which they deal. Expedients calculated to widen the gold points were extensively used by nearly all the Central Banks in the heyday of the Gold Standard between 1870 and 1914, but tended to disappear under the Gold Bullion Standards introduced after the war. In post-war years the possible range of variation set by the distance between the gold points was generally between ½ and 1 per cent. of the ruling rate of exchange. In 1928, for example, the dollar-sterling gold points were about ¾ per cent. apart and the franc-sterling gold

[1] If the margin is twopence in Great Britain and four cents in the United States, the gold export point from Great Britain will be 140s. 1d. — $34.98, and the gold import point 139s. 11d. — $35.02. The spread between the gold points due to the second factor will then be approximately ¼ per cent.

points about ½ per cent. apart.[1] The distance between the gold points was never quite constant, since the cost of transport of gold fluctuated with freight, insurance, and interest rates, and was gradually being reduced through the development of air transport, which, by saving time in transit, cut down loss of interest on gold shipments.

The Gold Standard provides a common medium of exchange, with which, in the last resort, international payments can be made. But only in the last resort. Importers do not normally send gold in payment for the goods which they import; nor do exporters fetch the same gold back again in payment for the goods which they export. Instead, importers procure from exporters the foreign exchange which they require and so save the double cost of transport of gold.[2] So long as the balance of payments is in equilibrium, and total exports are equal to total imports, there is no need for gold to move at all. It is only when imports and exports get out of line with one another that some residual transfer of gold is necessary in order to balance the two. Gold will flow in when importers have more foreign exchange than exporters will buy (so that its price drops to gold import point) and will flow out when the supply of foreign exchange is insufficient to meet the demand (so that its price rises to gold export point). The gold is purchased by, or obtained from the Central Bank which, since it is its duty to keep the price of gold constant, must hold a substantial stock of it, adding to the stock or drawing on it whenever the balance of payments is favourable or unfavourable.

If a country is on the gold standard the Central Bank must do more than *hold* a stock of gold and relieve any pressure on the exchanges by appropriate purchases and sales. It must also be able to *safeguard* its stock by taking action to check any drain of gold. To some extent it can rely on the flow of gold to reverse itself automatically, but too large a flow one way or the other may call for active intervention. Let us see how far the gold standard works automatically and what weapons can be used by the Central Bank to protect its gold stock.

**(a) The Gold Standard as an Automatic Standard.**—Once gold begins to move out to foreign countries, forces are automatically brought into play which tend to reverse the flow.

(i) The persons who export the gold pay for it with cheques drawn on their banks. The Bank of England, on receipt of these cheques, debits them against the banker's balances which it holds. Thus the

[1] In the earlier, pre-war, gold standard there seems to have been more "play" in the exchanges, the range of variation being 2 per cent. or even more. The greater rigidity of the post-war gold standard contributed to its ultimate downfall.

[2] A similar economy was made in domestic trade by the use of the inland bill (above, page 304). Exporters from any district sold the bills with which they were paid to importers who had remittances to make, so that the cost of moving gold to and fro was saved. Since gold was moved less frequently, the inland bill also dealt a blow at the highwayman, just as the foreign bill dealt a blow at the pirate.

joint-stock banks find on the one hand that deposits have been withdrawn in order to pay for the exports of gold, and on the other, that their cash reserves at the Bank of England have been reduced by an equal amount. Their reserve ratio falls automatically.

(ii) In order to restore their reserve ratio, the joint-stock banks are forced to call in loans, first from billbrokers and later from industry. Deposits with the banks fall as credit is restricted.[1]

(iii) The first effect of credit restriction is to raise interest rates. There is a fall in deposits, both because of withdrawals by gold exporters (the " primary " reduction) and because of credit restrictions (the " secondary " reduction); money, therefore, becomes scarcer, or " tight," and interest rates rise. The earliest and largest increase is in discount rates, since it is the money market which feels the immediate impact of credit restriction. The higher rates of discount raise the return on money invested in Britain relatively to the return elsewhere. There is an increased demand for sterling funds in order to take advantage of the higher return on them; and a reduced demand for foreign exchange, because the holding of foreign balances has become relatively less attractive. The increase in the demand for pounds and the reduction in the demand for foreign exchange combine to move the rate of exchange away from gold export point and the outward flow of gold comes to a stop.

(iv) It is not enough, however, that London should become a remunerative centre for the investment of money. The balances which are attracted from abroad, and which take the place of gold exports in the balance of payments, are mere stop gaps. So long as foreign balances continue to be attracted to London,[2] imports can remain in excess of exports without any movements of gold. But immediately no further balances arrive, gold exports have to be resumed. And if, for any reason, the balances are again withdrawn—as is so apt to happen with " hot " money—the evil is a hundred times worse for having been so long concealed. There is no real cure for an unfavourable balance of payments, but a fall in imports or a rise in exports. Nothing short of this can finally put an end to a drain of gold. If therefore, the gold standard is automatic in its working, the loss of gold must do more than raise interest rates and attract foreign balances; it must also limit imports, expand exports, and turn an unfavourable balance into a favourable one.

[1] *cf.* above, pages 329-30. The sale of gold by the Bank of England initiates a contraction of credit in exactly the same way as the sale of government securities.

[2] There is, of course, no *physical* movement of foreign money into London. All that happens is that foreigners buy pounds and sell foreign money in exchange: this foreign money is then used in payment for excess imports. In the same way, when we talk of money leaving the country, no money (except gold) moves: it is its ownership, not its situation, that changes. Some people sell pounds (which other people buy) in return for foreign moneys.

This *may* happen; indeed, it is certain to happen if we are willing to pay the high price which it involves. First of all the restriction of credit to industry reduces output and creates unemployment. The rise in interest rates works in the same direction. It alarms employers and discourages them from producing as large an output as usual, and at the same time, by raising the price of loans, it reduces capital outlay. Industry heads downwards into depression. At the outset, there is an immediate improvement in the balance of payments. Stock of goods, which cannot find a market at home are shipped abroad, imports of raw materials fall off, and less is spent out of the reduced incomes of the public (and especially the unemployed) on imported foodstuffs and manufactured goods. The later repercussions on the balance of payments are even more favourable. The pressure of unemployment forces the workers to consent to reductions in wages; costs and prices fall; Britain becomes a cheap market from which to buy and a difficult market in which to compete; exports rise, imports fall, and with the balance of payments favourable again, gold flows in and the whole downward movement is reversed.

The fall in wages, however, may be long in coming. The country may remain in a state of chronic depression, its balance of payments in equilibrium only *because* it is depressed. By impoverishing itself sufficiently, the country can be sure of reducing imports to match the fall in exports; it need only restrict credit, employment and purchasing power. But impoverishment, even for the sake of remaining on the Gold Standard, is not popular. The alternative of cutting wages[1] and clearing the way for credit expansion may be impracticable because of Trade Union opposition. The only escape from such a dilemma is to abandon the Gold Standard altogether—the course, which Britain in circumstance similar to those which we have been describing did in fact adopt in 1931.

So long as a country remains on the Gold Standard, therefore disequilibrium in its balance of payments is corrected in three more or less distinct stages. In the first, high interest rates attract balances from abroad. In the second, industrial depression reduces imports and may give a temporary stimulus to exports. And in the third, a fall in money wages clears away the depression and restores full equilibrium. If the original disturbance in the exchanges is favourable to the country, there will be three similar stages in which the transfer of balances to other countries is followed by increased employment and profits and later by higher money-wages. In either case the burden of final adjustment to the original disturbance is thrown on wage rates. This is, indeed, the gravest possible criticism that we can bring against the gold standard; it pre-supposes a plasticity

[1] Since prices would also fall, the cuts in money wages would not necessarily be cuts in real wages. The essence of the problem is to get prices down without simultaneously reducing real incomes.

of wage rates which does not exist and probably never has existed.[1]

**(b) The Gold Standard as a Managed Standard.**—The loss of gold from the Central Bank, therefore, tends to be automatically arrested and reversed by the forces which it releases. But these forces—of which credit restriction is the chief—may be slow in gathering strength, while the repercussions on the Bank's gold reserve are immediate and serious.[2] If the gold reserve is already small, and little in excess of the minimum legal requirements of the Bank, there is a real danger of panic. Foreigners, afraid that the country may be forced off the Gold Standard, and that gold or foreign exchange will then be obtainable only at a much higher price, rush to withdraw their balances, and by increasing the pressure on the Bank's gold stock, add to the very danger which alarms them. The Bank, therefore, may be forced to take drastic action in self-defence and to speed up or supplement the automatic correctives which are at work. Alternatively, if it has reason to believe that the loss of gold is temporary and that no tightening of credit is necessary in order to secure its return, it may seek to counteract and alleviate the direct impact on the money-supply of the loss of gold. In other words, the Central Bank may intervene in support of some definite line of policy; it may seek to *manage* the Gold Standard.

### The Technique of Management

**(1) Open-Market Operations.**—Of the weapons at the disposal of the Bank of England in managing the currency, either on the Gold Standard or off, open market operations rank first. By selling Government securities the Bank can initiate or reinforce a contraction of credit, and by buying Government securities it can promote an expansion of credit. *Whatever* assets the Bank acquires—government securities, gold, Treasury Bills—it creates deposits with itself (generally in favour of the commercial banks), broadens the base of credit, and

[1] The Gold Standard, it may be observed, worked much more smoothly during the period 1896—1914 when prices were rising and most variations in money-wages were *upwards*, than in the preceding period when prices and wages were falling.

[2] It seems to have been believed, when the 1844 Act was passed, that the loss of gold would not seriously deplete the central reserve held by the Bank of England: on the grounds that each £1 of gold exported would diminish the note circulation by £1 and so leave the Bank's gold reserve unchanged. As was pointed out at the time, however, the public's requirements of notes were inelastic, and responded later and indirectly to the loss of gold, not immediately and automatically. The loss of gold reduced deposits, not notes: yet it was to notes, not deposits, that the gold reserve was linked by the Act. As the "free" gold held by the Bank was rarely more than £10 millions to £20 millions, any substantial withdrawal required strenuous measures to counteract it, and there were real panics on several occasions (e.g. in 1847, 1857 and 1866).

gives the whole monetary system an inflationary impetus; whatever assets it sells, and whatever loans it calls in, it cancels deposits with itself, narrows the base of credit, and applies deflationary pressure to the commercial banks. The purchase of one asset is equal and opposite in its effects to the sale of another. A loss of gold, therefore, can be offset by the purchase of government securities; an inflow of gold can be offset (or sterilised) by the sale of government securities. By keeping its total assets constant, the Bank of England can insulate the whole credit system from the effects of gold movements; but only so long as it still has gold to sell when the flow is outwards or government securities to sell when the flow is inwards.

(2) **Bank Rate.**—Secondly, Bank Rate may be put up. Since Bank Rate is normally above the market rate and not an effective rate of discount for more than a very few bills, the rise in Bank Rate might seem powerless to affect the market. But by custom, changes in Bank Rate are the signal for changes in many other rates, including the call money rate at which billbrokers borrow from the joint stock banks. If the billbrokers have to pay more, they also charge more, so that the market rate of discount rises with Bank Rate. If market rate does not rise sufficiently, and remains well below Bank Rate, the Bank of England can soon make its rate effective by sales of bills in the open market. Usually, however, such sales are not necessary. The charge in Bank Rate is accepted as a pointer to the credit policy which is likely to be pursued and the money market raises its rates in anticipation of that policy.

(3) **Official and Unofficial Pressure.**—The Bank is in close touch with the leading personalities in the money market—some of whom are directors of the Bank—and it can influence rates by explaining its wishes to them. It can also, through its association with the Treasury, employ official pressure when necessary. For example, if the withdrawal of funds raised by new issues on the London capital market is weakening sterling, the Treasury may be induced to put an embargo on issues for foreign countries or to permit issues only on strict conditions.

**International Management.**

An essential part of good monetary management is co-operation with other Central Banks. Such co-operation helps to prevent unnecessary embarrassment to any country's banking system through excessive withdrawals of gold or of short-term balances. If, for example, one country has an adverse balance of payments which it is doing its best to correct, its task is made easier if foreign central banks adopt a cheap money policy, or if they offer large credits to tide over the period of adjustment. International monetary co-operation may also help to stabilise the general level of world prices and the volume of world production and trade. If each country tries to protect itself from the effects of a world depression by a beggar-my-neighbour policy of trade and credit restriction, each draws the

others into the depression like so many railway wagons during shunting operations. But if common action is taken to expand credit simultaneously in each country, the worst effects of the depression may be avoided.

**The Limits to Management.**

The Gold Standard, then, involves management; the preservation of a fixed par of exchange leaves scope for a choice of methods in which discretion and judgment are indispensable. The Central Bank may decide to offset a movement of gold which it judges to be temporary; it may try to reconcile the credit requirements of industry with the state of the foreign exchanges; and it may attempt in conjunction with other Central Banks, to gain control over world prices and trade. But management of this kind is possible only within narrow limits If the balance of trade is persistently unfavourable, no amount of management can prevent deflation. The alternative is either to devaluate or to leave the Gold Standard altogether.

**Devaluation.**

Devaluation means the adoption of a new and lower parity of exchange; the price of gold in terms of the currency is increased and the value of the currency, therefore, in terms either of gold or of other currencies is reduced. The effect of devaluation is, as a rule, to wipe out the unfavourable balance and to convert it ultimately into a favourable balance. Suppose, for example, that the pound falls in value from $5 to $4. Then British exports to the United States which were previously sold at $5 can now be sold at $4 without any reduction in the number of pounds received by the exporter. By quoting the same price in terms of sterling, therefore, British exporters will be able to undersell their foreign competitors; this will give a fillip to exports and reduce the unfavourable trade balance. On the other hand, if exporters quote the same price in terms of dollars, they will still sell as large a quantity of goods as before and the sterling value of exports will again increase. Similarly, imports into Britain will require to be sold at prices 25 per cent. higher if the proceeds in terms of dollars are to be maintained; for each pound will now fetch only four instead of five dollars. The rise in the price of imports will cause a reduction in the volume of purchases and also, if the demand is elastic, a reduction in the value of imports. This will reinforce the improvement in the balance of trade and relieve the pressure against sterling. When imports consist very largely of foodstuffs and raw materials, however, they will be in inelastic demand and devaluation, far from reducing their value, will actually increase it. In such circumstances devaluation may be comparatively ineffective as a means of correcting an unfavourable balance; it is conceivable even that devaluation might aggravate the unfavourable balance. In practice, a country's demand for imports and the foreign demand for its exports rarely fail to be elastic enough to produce a contraction in imports and an expansion in exports. Devaluation turns the

balance of payments in its favour and relieves the pressure on its exchanges.

Devaluation can hardly be a recurring expedient. If a country devaluates its currency every time it suffers from an unfavourable balance, it will destroy the confidence of traders and speculators in the fixity of its exchange rate and lead them to expect further depreciation. Without that confidence, half the value of the Gold Standard in facilitating international trade and investment is lost; the country has moved half-way from a system of fixed to a system of free exchanges.

If a country decides against both deflation and devaluation, it must abandon the Gold Standard altogether. It may continue to observe most of the rules of the Gold Standard game by setting up an Exchange Equalisation Fund, or it may leave the exchanges to be regulated by market forces, or it may institute various measures of exchange control.

**The Exchange Equalisation Account.**

The first of these devices was used in Britain after 1932 to control fluctuations in sterling. A small committee of the Treasury was set up to manage an Exchange Equalisation Account with a view to ironing out erratic fluctuations in exchange rates. The Account was empowered to borrow £150 million through the issue of Treasury Bills and to use the resources so obtained to purchase gold and foreign exchange. Later the borrowing powers of the Account were increased, first to £350 million and then to £550 million. By exchanging sterling for foreign exchange or vice versa, the Account is in a position to keep the sterling exchange rate constant until it has exhausted its supply either of sterling or of foreign exchange. The speculator who is reluctant to hold francs can buy sterling from the Account, which generally converts its purchase of francs into gold; later, when the speculator repatriates his capital—that is, re-converts it into francs—the transaction is reversed, the Account selling gold (or francs) and buying sterling. Thus the Account can soak up any "hot money" which is temporarily deposited in London and release it whenever the depositor wishes to withdraw his balance. When there is a persistent movement against sterling, however, the Account is powerless to control it, since its holdings of foreign exchange are limited. Its object is to prevent fluctuations which are due to speculation and to keep them from affecting our trade and credit, not to prevent fluctuations which are due to "real" causes operating through the balance of payments. In order to increase its powers against speculators, the operations of the Account are kept secret, since unknown resources are more formidable than known. At first this secrecy was carried to extremes, especially as fairly accurate guesses could be made by economists.[1] In 1937, however, half-yearly statements began to appear at an interval three months after the dates to which they applied. During the war these statements were suspended.

[1] Notably by Mr. F. W. Paish: ("The British Exchange Equalisation Fund," *Economica*, 1935).

The Exchange Equalisation Account partially reproduces Gold Standard conditions. For example it limits the range of fluctuations in exchange rates. In each year after 1933 until the crisis of 1938 the margin between the highest and lowest price of dollars steadily diminished. But the Account retained discretion, as the Bank of England did not between 1925 and 1931, to allow the exchange rate to depreciate or appreciate. There was more elasticity, and consequently more uncertainty, in the control of exchange rates. There was also one curious difference in the effect on interest rates of movements of foreign balances. Under gold standard conditions an inflow of balances from abroad necessarily reduces interest rates since it is accompanied by an inflow of gold and an expansion in the money supply. But if sterling balances are purchased from the Exchange Equalisation Account in exchange for sales of francs or dollars, rates of interest may *rise*. For the Account will be forced to sell Treasury Bills in order to obtain the necessary sterling funds, and if the sellers of francs or dollars wish to hold liquid balances rather than bills, there will be no corresponding increase in the demand for bills. Such a situation, however, was rare and never of really great importance. Any persistent transfer of funds from or to Great Britain generally led sooner or later to a transfer of gold from or to the Bank of England, the Exchange Equalisation Account acting as a kind of safety valve and passing on surplus gold in return for sterling funds, or buying gold from the Bank whenever its supply of foreign exchange was running low. In the long run, the antithesis with gold standard conditions tended to disappear. Only the difference in elasticity remained.

**Forward Exchange.**

When there is no fixed parity, traders are exposed to a good deal of uncertainty. If they sell for delivery in three months' time, they cannot know what value the foreign currencies in which they are paid will have at the end of that time. The technique of the forward exchange market can be used, however, to transfer their risk of fluctuations in exchange rates to some middleman, e.g., the banks. The exporter will be paid in pounds on the basis of the current rate of exchange, and will sell his right to receive foreign exchange three months hence. The bank will be able to sell these rights to importers who have to make payments in three months' time. If the banks are left with some forward exchange still on their hands, they can " hedge " by selling that exchange " spot," borrowing abroad what they sell, and paying off the loan at the end of three months. There will be two rates of exchange, the " spot " rate for cash transactions and the " forward " rate for " future " transactions, i.e., transactions in titles to foreign currencies at some future date. The margin between the two rates will depend upon the cost of borrowing in the two centres and upon the uncertainty of the future course of the exchanges.

**Free Exchanges.**

When the rate of exchange is free to vary, it will be governed, as we have seen, by the balance of payments. The Central Bank may

still employ its holdings of gold and foreign exchange in order to damp down violent movements in exchange rates but such action is no longer legally enjoined upon it. It will be, in the main, the ordinary forces of the exchange market—the pressures to buy and sell foreign exchange—that govern exchange rates.

Suppose now that we are wondering whether a country's exchange rates are *normal* or in equilibrium. How are we to decide? The answer is that we must study the country's balance of short-term indebtedness. If the country is losing gold or is selling foreign balances or is being forced to borrow on short-term in order to finance its imports, then its exchange is overvalued and is likely to depreciate. But if the country is gaining gold or accumulating large balances of foreign exchange, because it is not importing on the same scale as it is exporting, then its currency is undervalued and is likely to appreciate. It is not necessarily true, however, that the equilibrium rate of exchange is one at which no change in the country's short-term indebtedness will be taking place. The country may be a centre in which there is an increasing and natural tendency for the short-term funds of other countries to find investment. Or the country may be trying to make a gradual addition to its holdings of gold and foreign exchange in order to safeguard itself against the effects of violent fluctuations in its exports.

**Purchasing Power Parity.**

Some economists have suggested a rather different test of equilibrium in the rate of exchange. They have argued that exchange rates tend to be pulled into line with what are called "purchasing power parities."[1] Sometimes this is taken to mean that if £1 has the same purchasing power in Britain as $4 in America, then the equilibrium rate of exchange is £1—$4. Sometimes with more justification, the theory is modified so as to refer, not to a comparison between the purchasing power of the pound and the dollar at any particular time, but to a comparison between the *changes* in their purchasing power since some base date. If, for example, the purchasing power of the pound has been halved while the purchasing power of the dollar has been doubled, then, it is argued, the dollar will tend to exchange for four times as many pounds as before. In other words, changes in the external value of a currency (its value in exchange for other currencies) tend to reflect changes in its internal value (its purchasing power over commodities).

The first version of the theory is clearly untenable. Every tourist knows that it is cheaper to live in some countries than in others and that, since prices are higher in America than in Britain, he can buy more with one pound than with the $4 for which it exchanges. Some prices, the prices of traded goods, must be more or less in line in

[1] The "purchasing power parity" theory is generally associated with the name of Professor Gustav Cassel, but the gist of the theory was put forward a century earlier during the controversies over the Bullion Report of 1810.

both countries; but other prices—fares, hotel charges, lawyers' fees, etc.—need not correspond because there is no way by which people living in one country can take advantage of cheap services provided in another. The second version of the theory is more reasonable. It is clear that if prices rise in any country (so that its currency loses purchasing power) there will be a strong tendency for exports to diminish and for imports to increase; the balance of payments will become unfavourable, and exchange depreciation will result. There is a presumption, therefore, that changes in purchasing power will be followed by similar changes in exchange rates. But this presumption must not be interpreted too rigorously; the changes in purchasing power and in exchange rates need not be in exact accordance with one another.

If we take the price of any staple traded commodity, such as wheat, we are likely to find that, converting at current rates of exchange, this price is the same in all countries; allowance being made for transport costs and tariff barriers. So long as the rate of exchange is \$4—£1, wheat selling at a dollar a bushel in Chicago will sell at rather more than 5s. a bushel in Liverpool. If the Liverpool price rises, more wheat will be shipped until the local scarcity disappears. In the same way a fall in the Liverpool price will reduce imports until parity with the Chicago price is re-established. Or if the rate of exchange depreciates to \$3—£1, the Liverpool price—failing any change in price in Chicago—will be marked up automatically by 1s. 8d. a bushel. What is true of wheat is true also of other traded goods. Prices in each country must conform to one another at current exchange rates whatever these rates happen to be. Thus it is useless to compare changes in the purchasing powers of two currencies over *traded* goods as a test of equilibrium. For unless freights or tariffs have altered,[1] these changes in purchasing power *must* be equal whether exchange rates are in equilibrium or not. It is the prices of traded goods which are pulled into line with exchange rates rather than exchange rates which are pulled into line with the prices of traded goods.

When we turn to domestic goods—that is, goods which do not enter into international trade, such as house-room, milk, services of all kinds—we find nothing to couple their prices in one country rapidly and automatically with prices in other countries. If, for example, domestic goods double in price in Britain and remain unchanged in the United States there is no obvious reason why the sterling-dollar exchange rate should immediately fall by 50 per cent. Purchasing power parities measured in terms of domestic goods, therefore, may certainly diverge from exchange parities. But can the divergence

[1] A rise in British tariffs or in freight charges on British imports will increase the margin between the price of traded goods in Britain and abroad. Prices will tend to rise in Britain, but since purchases from, and payments to, foreigners will fall, the balance of payments will be more favourable and the rate of exchange, instead of depreciating will probably appreciate. Thus purchasing power and exchange parities will diverge from one another.

*continue* ? Is the divergence a measure of disequilibrium in exchange rates and a safe basis on which to forecast their future course ?

It should be noticed, first of all, that the divergence can equally well be expressed as a divergence between the price levels of domestic and traded goods. The link between exchange-rates and the price-level of traded goods must also bind the general price-level unless domestic goods fail to keep in line with traded goods. If the price level of domestic goods lags behind the price level of traded goods, the purchasing power of the currency will also lag behind movements in exchange rates. And if equilibrium requires an equal change in purchasing power and in exchange rates, it must also require, therefore, an equal change in the price levels of domestic and of trade goods.

Now there are strong forces linking together these two price levels. If either group of commodities becomes dearer, the most likely cause is a rise in the cost of wages and materials which, if it is at all great, is fairly certain to spread to the other group of commodities. A persistent rise in wages in the export industries, for example, will soon lead to a sympathetic rise in wages in allied industries supplying the home market. First of all, domestic workers in the same craft will share in the rise, then workers in the same grade of labour, and finally workers of greater or inferior skill. Similarly, with raw materials. A rise in the price of steel affects ship plates for export equally with girders for offices and shops at home. Prices are stitched by competition into a fabric which resists distortion. Pressure at one point is felt on all sides, and, although the fabric may give for a time, the old price relationships are soon restored.

It is to this inertia of the price structure that we must turn if we wish to justify the purchasing power parity theory. If the prices of traded and non-traded goods keep in step, so also will exchange rates and purchasing power parities. But we must be careful not to exaggerate the stability of the price structure. Dissimilar changes in the level of output, the unequal incidence of invention, the immobility of labour, and the imperfections of competition make for irregular movements in wages and in prices in different industries, and even in large groups of industries. The experience of the so-called " sheltered " industries —those which supply the domestic market and are immune from foreign competition—need not parallel the experience of the " unsheltered " industries—those which produce traded goods and are, therefore, in competition with foreign industries. Indeed, it is not unusual for the fortunes of these industries to be widely different. Now if, as may happen, costs in the sheltered industries increase, while costs in the unsheltered industries do not, the country's competing power *vis-à-vis* other countries will be practically unchanged and exchange rates will not be much affected. But the level of prices will be higher and the purchasing power of the currency will be correspondingly lower. The test of purchasing power parity will have broken down.

A similar breakdown takes place whenever there is a change in a country's terms of trade. Suppose, for example, that of two countries

each consumes equal amounts of its own and of the other country's goods. A fall in the price of one country's exports will then affect prices (but not incomes) similarly in both countries, and will at the same time cause a depreciation in the rate of exchange parallel to the fall in export prices. That is, the rate of exchange will alter without any similar alteration in purchasing power parity.

A divergence between the rate of exchange and purchasing power parity may also arise because of international capital movements. If a country is lending large sums abroad, or if speculators are transferring balances to other financial centres, heavy sales will be made of the country's currency and its value will depreciate. But prices—at any rate, domestic prices—will not be directly affected by these transfers of capital, and the depreciation in exchange rates will not be accompanied by an equal fall in the purchasing power of the currency.

Once the appropriate qualifications have been made, the purchasing power parity theory loses much of its precision. Its popularity after the last war was due to the emphasis which it laid on inflation as the source of exchange depreciation. It suggested that the proper way to bring that depreciation to a halt was the restriction of credit inside the country. This would stop the rise in prices and, therefore, also the depreciation. Beyond this rather elementary piece of analysis, however, the theory has seldom penetrated. It is an unsatisfactory short-period theory because it makes no provision for the influence of speculation, capital movements, and so on; and it is an unsatisfactory long-period theory, because it fails to take account of the fact that wages and prices in different groups of industries do not move together.

**Fixed versus Fluctuating Exchanges.**

It is common to contrast the Gold Standard with a system of free exchanges and to ask which is preferable. But this antithesis is rather unreal. There must always be some distance between the gold points, however small, so that, even on the Gold Standard, the exchanges are bound to fluctuate a little. On the other hand, the device of the Exchange Equalisation Fund can secure some stability in the exchanges even when there is no *de jure* Gold Standard. It is true that countries which are on the Gold Standard are supposed to set known and certain limits to the fluctuations of their currencies, and that abandonment of the Gold Standard removes these limits. But devaluation has taken place so frequently that the limits are not in fact known with complete certainty. Our problem is not so much *whether* exchanges should be allowed to fluctuate, but *how much* they should be allowed to fluctuate.

The chief advantages of the Gold Standard are twofold. First, it offers great convenience to foreign trade and investment. By keeping the price of foreign exchange steady it helps to eliminate uncertainty from transactions with foreigners, and so assists the free flow of goods and of capital across national frontiers. It links

national markets together into a single world market which, by its greater width, gives encouragement to division of labour on an international scale. We are able to buy and borrow abroad almost as easily as at home, whereas with fluctuating exchanges we have to face the contingency of loss if our calculations are upset by depreciation or appreciation of exchange rates. Abandonment of the Gold Standard, therefore, must mean one more hindrance to international trade and investment, although such hindrances would seem, as a rule, to bring only a doubtful and transient advantage to one country and to impoverish the others.

To this it can be replied that the mobility of goods and of capital which the Gold Standard promises can be bought at too high a cost. Each country is forced to maintain more or less similar rates of interest, since any considerable margin between rates in different financial centres will cause a movement of funds very dangerous to existing exchange rates. Each country, therefore, is forced to deflate and (less cogently) to inflate when other countries are deflating or inflating. It must bring its price level into line with world (gold) prices. Freedom to depreciate when necessary means freedom to reduce interest rates and to attempt to expand employment. Admittedly, if all countries possess such freedom the advantage to each may be greatly diminished, since each can meet depreciation with depreciation, and uncertainty may be increased all round. Nevertheless depreciation does offer countries which have difficulty in preserving equilibrium in their balance of payments a weapon with the help of which wage cuts can be avoided and unemployment (for a time at least) reduced.

The controversy generally turns on the importance which is attached to international trade and to the resumption of foreign investment on a large scale. Those who look to an expansion of the home market to absorb unemployment, and consider international trade of secondary importance, are not likely to be enthusiastic about an international standard. But those who look to a revival of foreign investment as a possible stimulus to our export industries, and foresee some difficulty in transferring labour from the depressed (export) industries to the expanding (home) industries, generally support a return to a (possibly modified) Gold Standard.

## CHAPTER 28

## UNEMPLOYMENT AND THE TRADE CYCLE

*. "There have been such unheard-of prices and unchristian rates of interest and overcharges in business and commerce as never before in this world . . . so that many citizens through their rashness, pride and arrogance in letting themselves appear to grow great with foreign money and carry on big business . . . to their great damage did not foresee that the interest would eat up everything, and eat them up with it, until ruin had fallen upon them."*

BURGOMASTER OF LÜBECK, 1620.

*"A boom is a situation in which over-optimism triumphs over a rate of interest which, in a cooler light, would be seen to be excessive."*

J. M. (*now* LORD) KEYNES, 1935.

## VOLUNTARY UNEMPLOYMENT

THERE are still people who regard unemployment as a moral problem, originating in personal failings and curable only by personal effort. The man who fails to find work is presumed to be lacking in backbone and a proper spirit of independence. There is a job for him somewhere if he will look for it persistently enough ; others are successful, so why not he ? If we have more unemployment nowadays than in the nineteenth century, it is because we have made unemployment so much more attractive. Remove all temptation to idleness and unemployment will cure itself. So extreme a view as this can be held only by persons who are blind to the facts. It pre-supposes that there is a large standing army of permanently unemployed workers who, rather than earn an honest living, deliberately reject offers of employment and make no real effort to find work for themselves. But there is no such "standing army." The vast majority of those who are unemployed on any given date have been in work at some time within the preceding year, and will be in work again at some time in the succeeding year. At least half of the insured workers in Great Britain have been out of work at some time or other ; and if some are unemployed more frequently and for longer periods than others, this is not because of their "moral" failings, but because, as marginal workers, they are the first to lose their jobs in bad times and the last to be taken on again when trade revives. We cannot suppose, when we find three million workers unemployed at one time and only one million at another, that two million workers have suddenly overcome their laziness and experienced a conversion to the call of work. The more natural conclusion is that finding employment is like a game of musical chairs in which some of the chairs are from time to time removed and later replaced. It is not love of idleness but lack of opportunity to work which is the main cause of unemployment. We must look for an explanation not to the supply of labour and the motives of workers, but to the demand for labour and the motives of employers.

Nevertheless, although it is a travesty to suggest that unemployment is the fault of the unemployed, *some* unemployment is voluntary. There are, and have always been, social parasites in all classes who choose to live on their fellows rather than bear their share of the burden of work. It is true, too, that continuous unemployment may so paralyse a man's spirit, and so ruin his skill (or even his physique), that he has neither the inclination to look for a job nor the ability, when placed in one, to hold it. He gets into a routine of unemployed life which he is afraid to change for fear of having to learn it all over again—as anxious to avoid a temporary spell of employment as he was formerly anxious to put an end to a temporary spell of unemploy-

ment.[1] Again, the system by which unemployed workers in Great Britain are given family allowances, while employed workers are not, often puts a premium on idleness since a married man with a large family may actually find himself out of pocket if he accepts employment and is frequently only a few shillings to the good for a hard week's work. These facts cannot be disregarded. But there is no reason to believe that voluntary unemployment accounts for more than a small proportion of total unemployment. The larger problem is one of the involuntary unemployment of workers who, however willing, are unable to find work at current wage rates.

### Involuntary Unemployment

**(1) Normal or Residual Unemployment.**—Even in war-time, when the demand for labour is intense, unemployment never entirely disappears. A residue remains which, in more normal circumstances, may reach quite considerable dimensions. This residue includes, first of all, seasonal unemployment. There are few industries in which it is possible to maintain steady employment all the year round; either there are seasonal peaks in demand (e.g., in the manufacture of umbrellas, toys and motor-cars), or weather conditions hold up work at certain seasons of the year (e.g., in agriculture and building). These seasonal irregularities do not necessarily create unemployment. Sometimes it is possible, without excessive inconvenience, to make use of the services of married women in peak periods, or to provide off-peak employment for regular workers by dovetailing one kind of work with another (e.g., by making straw hats and bowler hats in the same works). Some workers, too—the salaried staff, for example—may be able to cope with seasonal pressure by working harder, slacking off again once the rush is over—a practice of particular importance in peasant communities where the problem of seasonal unemployment tends to be transformed into one of *under*-employment during the winter months. But in spite of all these alleviations, seasonal factors do generally create actual unemployment. Each industry tends to keep a permanent reserve of labour for which it finds employment only at certain times of year.

Unemployment also arises in industries in which demand fluctuates erratically. A dock labourer, for example, may work at high pressure for a few hours or a few days when a large number of ships put in, and find himself unemployed for the rest of the week when the berths are empty. Anyone who is engaged in repair work, or in making goods influenced by fashion, is likely to have a similar experience. Bad organisation, accident and failure of supplies may also lead to unemployment. For example, plaster work scheduled to be done on a housing scheme may be held up, leaving the plasterers idle, because the bricklayers are behindhand, or because of a shortage of timber, or because a heavy fall of snow has held up transport.

[1] For a full study of long-term unemployment, see "Men Without Work (A Report to the Pilgrim Trust)".

It is common, too, for some days or even weeks of unemployment to elapse between the conclusion of one job and the discovery of another. There may be vacancies in plenty in other places and industries, but none of them, given human immobility, quite suitable. However well-organised the labour market, therefore, and however keen the demand for labour, unemployment will persist. Exactly how much will persist and is really unavoidable, it is not easy to say; but we may guess that, except in circumstances which could not be indefinitely prolonged, the percentage of workers unemployed would be unlikely to fall below two.

(2) **Structural Unemployment.**—Hardly less avoidable than residual unemployment, and generally classified with it, is unemployment caused by changes in the structure of industry: by changes in industrial organisation and technique, by the decline of major industries, and by the migration of industry from one region to another. All of these played an important part in the post-war unemployment problem which Britain and other belligerent countries had to face in the 'twenties and 'thirties. Technique advanced at a prodigious pace, displacing workers who could not always be re-absorbed in other industries. Schemes of industrial reconstruction, designed to eliminate surplus capacity and to concentrate output in fewer and larger plants, led to wholesale dismissals from the works which were shut down. Important war industries, such as mining, iron and steel, and ship-building, had to make a swift change-over to a peace-time basis in competition with the over-expanded war industries of other countries. The export industries, cut off from their markets by four years of war, found that native industries had grown up there in the meantime, and showed no signs, thanks to high tariff barriers, of disappearing. Above all, the new and growing industries, such as motor-car manufacturing, instead of expanding alongside the older and declining industries and absorbing unemployed workers in the vicinity of their homes, grew up in the very parts of the country where there was least unemployment, and created a major problem of labour transference from contracting to expanding areas. No group of workers suffered more from these changes than the coalminers. Research contributed far more to fuel economy than to the discovery of new uses for coal; even when coal held its own against the competition of oil, electricity and gas, it was so carefully economised that requirements were often cut in half. At the same time, mechanical coal-cutting was increasing output per head and reducing labour requirements. In some areas the mines were becoming worked out, or were being forced out of production by high costs. In others, which were heavily dependent upon export markets, the loss of important markets to foreign coalfields was limiting output and employment. After a brief post-war boom, in the course of which thousands of additional workers were attracted to the mines, the industry plunged into a deep and prolonged depression. **Employment fell from 1,212,000 in June 1923, to 638,000 in June 1932,**

and since mobility was low in comparison with other industries, and opportunities of alternative employment rare, a serious problem of long-term unemployment came into existence in the mining areas. Of 325,000 workers who had been continuously unemployed for over a year in June, 1936, no less than 81,000 were miners, and many more must have had only occasional spells of employment during the year.

Structural unemployment of this kind represents the " hard core " of unemployment. It is not intermittent like other types of unemployment, but results from a permanent change in opportunities of employment to which workers have to adjust themselves. Men engaged in work of one kind or in one place find themselves shut out from their customary employment and forced to change their occupation or domicile. The more immobile they are, and the greater the structural change in progress, the greater is the resulting unemployment. In pre-war Britain the ageing of the population reduced mobility, while the changed balance between export and home markets, between capital and consumption goods, and between north and south, put a heavy strain on the adaptability of industry. Structural unemployment, therefore, was unusually high.

(3) **Cyclical Unemployment.**—The causes of unemployment which we have so far discussed have operated only within individual industries or groups of industries. We have seen no reason to expect regular oscillations between boom and slump, covering the whole field of industry so that all, or nearly all, industries are simultaneously prosperous or depressed. Yet we know that, in what is called the Trade Cycle, such regular oscillations do take place. For at least a hundred years, the demand for labour has fluctuated in waves or cycles, in a steady alternation between boom and slump, the length of the cycle varying from as little as six to as much as ten years. In periods when the long-term trend of prices has been downward—for example, between 1874 and 1896, and between 1920 and 1933—the cycle has lengthened and depressions have been long and severe; in periods of rising prices, fluctuations have been more rapid and less violent. Trade cycles have not in the past, and do not even now, synchronise perfectly in all countries; but differences in timing have steadily diminished, and a major depression in any of the leading industrial countries is soon communicated to the others.

The trade cycle is primarily an employment cycle. The main feature of a boom is a shortage of labour and the main feature of a slump is unemployment. But associated with these fluctuations in employment, partly as cause, partly as consequence, are fluctuations in a host of other things: in prices, output and incomes; in exports and imports, the birth-rate, budget results, drunkenness. A common cyclical pattern stamps itself on many things between which there is little direct connection—for example, the marriage-rate and the consumption of beer—suggesting causal relationships where none exist.

There are four more or less distinct phases in the typical cycle:

(*a*) *Revival.*—First, starting from the bottom of the depression, comes a period of revival in which there is a growing feeling of confidence and an atmosphere more congenial to enterprise. Larger orders are placed with manufacturers either in order to replenish traders' stocks or in order to renew or extend capital equipment. To fill these orders, manufacturers are forced to engage more workers, and the newly-employed workers spend more freely out of their increased earnings. Spending responds also to the increase in profits which manufacturers are able to earn by making fuller use of their plant. Additional orders are placed to provide for the increase in consumers' expenditure, and these orders give a fresh stimulus to production. As fast as more goods are despatched to shopkeepers, more income is generated and more goods are bought. Thus the upwards movement gathers momentum like a pendulum and gradually passes into the second phase.

(*b*) *Prosperity.*—As output nears the limits of industrial capacity, prices begin to rise. Raw materials rise fastest and earliest because the supply is highly inelastic in the short period, while the demand, in view of the comparative insignificance in total costs of many raw materials, is generally also inelastic. As time goes on, something of a scramble for raw materials may develop, since producers will wish to ensure fulfilment of contracts into which they have entered, and prices may then be driven up to a level which few expect to see maintained. The rise in prices is by no means confined to raw materials; it is common to all industries in which an increasing demand pushes against a relatively fixed supply. Whenever a "bottleneck" appears, and the scarcity of some key type of equipment, labour or raw material limits output, price is forced upwards. Confronted with a rising cost of living, and strengthened by the decline in unemployment which accompanies trade revival, wage-earners begin to claim wage-advances, and these, when granted, drive costs and prices still higher. Meanwhile investors and business men, in a spirit of incautious optimism, are sinking their capital in projects which discount the future much too generously by assuming the maintenance—and often the intensification—of boom conditions. Capital is invested in competing schemes, each begun in ignorance of the others, and each doomed to disappointment when the others come into production. The capacity of important industries is expanded with a view to profiting from abnormally high prices, or from an abnormally active demand; without regard to the depression in prices which the expansion in capacity will of itself produce. A boom sets in.

(*c*) *Crisis and Recession.*—The boom—if it is not allowed to develop into an inflationary spiral—breaks for a variety of reasons. The monetary authorities may take fright and restrict credit. Or, since rising prices tend to create an unfavourable balance of payments, the Central Bank may find its gold reserve falling to a dangerously low level and put up interest rates in self-defence. Or the bubble of

confidence may be pricked by the failure of a few rash ventures or by the disinclination of merchants and producers to place repeat orders at prices which they regard as excessive and abnormal. Or projects of capital construction may reach a temporary satiety. Whatever the reason, a point of crisis is reached and the whole upward movement is reversed. Enterprise is checked and orders are cut to a minimum. Prices begin to fall, and the expectation that the fall will continue checks buying. Output is reduced and with it, demand, in a vicious circle of action and reaction. Industry heads downwards into depression.

(*d*) *Depression.*—To prolong a peace-time boom—except by blowing it up into an inflation—is one of the most difficult tasks that economists can conceive of; to prolong a slump is one of the easiest. A boom almost inevitably comes to an end; but there is nothing evanescent about a slump—it may last for generations. In fact, however, there seem to be forces in the economic system which limit the severity of depressions; sooner or later a kind of automatic brake is put on and industry, instead of slumping endlessly, begins to bump along the bottom, waiting for some new impulse to carry it up again. Just as a boom contains the seeds of depression, so a slump seems to contain the seeds of revival. But what are these seeds? What is it that arrests and reverses an industrial depression? This is a question to which no satisfactory answer can be given until the *cause* of the trade cycle and of cyclical unemployment has been investigated.

**Theories of the Trade Cycle.**

A large number of theories have been put forward, many of which can be unhesitatingly rejected. It is absurd, for example, to explain unemployment in terms of over-production, for so long as many of our wants go unsatisfied a *general* over-production of goods is impossible. If there is a surplus of unsaleable goods—and still more, if there is a surplus of labour for which no employment can be found—it must be because there is a deficiency of effective demand, not because we have more than we want. If this deficiency were to disappear—as happens in times of boom—we should find ourselves producing too little, not too much. So long as we realise that it is the slump which is abnormal, and not the boom, we cannot attach much importance to over-production theories. There may, however, be genuine over-production of particular commodities and a corresponding under-production of other commodities. Output, that is, may be lop-sided, so that workers remain unemployed in some industries while there is a labour shortage in others. Such *structural* unemployment always seems larger in a depression than it really is. For workers do not move in search of employment when there is unemployment wherever they go; they move only when there are vacancies to move to. Thus as trade revives, mobility increases, and adjustments which were hopeless within the rigid framework of depression are made with comparative ease in the more expansive atmosphere of a boom.

Moreover, since it is generally the most depressed industries which recover most ground as trade improves, the need for adjustment is diminished simultaneously with the difficulties of making it; outputs which seemed lop-sided in the depression become normal in the ensuing boom.

Other suggested causes of unemployment are the introduction of machinery, a chronic deficiency of purchasing power, and the temperamental instability of business men. The first of these, in a period of rapid technical progress, may undoubtedly give rise to what is sometimes described as "technological unemployment." Workers who are displaced by machinery are not always re-absorbed into employment immediately. But unless there is a shortage of jobs for some *other* reason, most workers, if reasonably adaptable, will ultimately find employment. There is no reason to believe, looking round the industries which suffer most from unemployment, that mechanisation has played any very large part in causing it; indeed, some of the industries in which mechanisation has gone furthest, such as the motor industry, have simultaneously taken on large numbers of additional workers. It is a little old-fashioned, too, to think of technical progress in terms of mechanisation; more coal-miners were thrown out of work by new methods of utilising coal than by the introduction of mechanical coal-cutters.

If the first suggestion is inadequate, the second is simply fallacious. There may be a deficiency of purchasing power in times of depression, but there is no such deficiency in times of boom; there is no *chronic* deficiency of purchasing power such as Major Douglas, for example, alleges to exist.[1]

The third suggestion—fluctuations in business confidence—does not take us very far. When a crowd collects there is usually something for it to stare at; although the more people are there already, the easier it is to attract others. In the same way, when a mood of optimism grows, there are usually some real grounds for it, although once it begins to spread it quickly gathers force. The prevailing mood, whether it is one of optimism or of pessimism, tends to contribute to its own justification. If it is one of optimism, enterprise is encouraged, employment increases and the prosperity to which people looked forward comes into existence; if it is one of pessimism, enterprise is discouraged, unemployment increases and the community is plunged into the depression which was so widely anticipated. But these changes of mood are unlikely to be spontaneous over the whole field of industry. Behind them we are fairly sure to find some real change—possibly in the world of politics, but more usually in the profit situation—and it is in this change, and not in a change of mood, that upswing and downswing originate.

[1] For a fuller discussion of the Douglas theory, see above, pages 41-2. Of the many refutations of the theory published by economists, the best is that of Mr. Hugh Gaitskell in a chapter contributed to "What Everybody Wants to Know About Money," (edited by G. D. H. Cole).

**The Theories of Economists.**

None of the theories so far discussed—except possibly the last—has ever been held by any reputable economist. Economists themselves, however, have by no means reached agreement as to the cause of unemployment. One view, held by Mr. R. G. Hawtrey, is that the trade cycle is "a purely monetary phenomenon." The mainspring of cyclical change is Bank Rate. A reduction in Bank Rate—and in other short-term rates of interest—encourages merchants to carry larger stocks of goods: partly because the interest charges on the merchants' overdrafts (which are a high proportion of the total cost of carrying stocks) are simultaneously reduced; and partly because a reduction in Bank Rate is traditionally the forerunner of business revival. The increase in stocks is soon checked since, as we have seen, larger orders mean increased purchasing power and increased purchases. Trade continues to expand until the Central Bank, finding its gold reserve beginning to melt away, raises Bank Rate and forces merchants to curtail orders to manufacturers. A depression sets in and lasts until Bank Rate is brought down again. By raising Bank Rate, argues Mr. Hawtrey—and on this there is general agreement—we can put a stop to any boom; by reducing Bank Rate—and here many economists are sceptical—we can end all, or nearly all, depressions.

A second theory, which was developed by W. S. Jevons, lays stress on harvest fluctuations. There is some evidence of a cycle in weather conditions of approximately the same length as the trade cycle. The weather governs harvest yields, and harvest yields—since agriculture is much the largest industry in the world—react on industrial prosperity. Most economists are prepared to agree that harvest fluctuations play an important part in the trade cycle; but these fluctuations are now regarded as affecting the *timing* of revival and recession rather than as causing the cycle itself.

The views of Mr. J. A. Hobson may also be mentioned. Unemployment, says Mr. Hobson, is the result of chronic over-saving. Spending creates a demand for labour, but saving does not. If, therefore, we wish to cure unemployment, we must spend more freely. But because income is unequally distributed, some people have more money than they care to spend, while others, who would like to spend more, lack purchasing power. Until income is more equally distributed, the trade cycle will remain with us. What is of value in Mr. Hobson's view is incorporated in the theory put forward by Lord Keynes—a theory which, since it commands the widest approval, we must study in some detail.

**The Significance of Investment.**

We may start from a fact. Fluctuations in employment have always been much more pronounced in industries making capital goods than in industries making consumption goods. In a depression, unemployment is greatest in industries like shipbuilding and steel-making, and least in industries like transport and distribution.

Roughly speaking, the more durable the product, the higher the unemployment percentage. This strongly suggests that the clue to the unemployment problem lies somewhere in the process by which capital is accumulated.

Now there are two independent sets of decisions about capital accumulation which are taken by different people, and which need not synchronise with one another in such a way as to preserve full employment. One set of decisions concerns saving and spending, and is taken by consumers; the other set concerns investment or loan-expenditure,[1] and is taken by manufacturers, estate companies, local and central government authorities, etc. The first set of decisions involves the release of real resources, including labour, from meeting the immediate requirements of the savers; the second set of decisions involves the employment of real resources (including labour) on works of capital construction. There is no automatic machinery for marrying these decisions to one another so as to see to it that the resources released by the savers are absorbed by capital expenditure.

Habits of saving, as we have seen, are not subject to violent change; saving is generally a residue after customary needs have been met and varies mainly with income. Investment, however, is inherently unstable. This is so for a variety of reasons.

First, capital expenditure is *postponable*, whereas expenditure on goods for current consumption cannot easily be cut down or expanded. If no rubber trees were planted for a year or two, the effect on the rubber supply would be relatively trifling; but there would be a most serious shortage if no rubber trees were tapped for six months.

Secondly, the requirements for which capital expenditure provides are less calculable because more distant in time than current requirements. The longer the period in view the more difficult it is to predict changes in demand, in the cost of operating or renewing plant, in the rate of interest, in technique, and so on. The longer the period for which capital has to be sunk, therefore, the more responsive is capital expenditure to the changing expectations of business men about the future. A spirit of caution will bring about a sharp fall in investment; if the uncertainty passes, there will be an equally sharp rise in order to make up lost ground.

The same reasoning applies whether capital goods take a long time to manufacture or whether their length of useful life is so great that a long interval must elapse before their original cost is recovered. Railway-building in the nineteenth century proceeded by fits and starts partly because railways last such a long time. No one could hope to pass judgment with any degree of confidence on their profitability in fifty, or even twenty years' time; and the pace of construction tended

[1] It should be clearly understood that in this context "investment" means, not the purchase of a piece of paper (whether bond, stock, share, mortgage or the like), but capital expenditure on some instrument of production (e.g., the manufacture of machinery, the construction of roads, bridges, houses, etc., afforestation and so forth).

to fluctuate, therefore, with the whims and moods of investors, played upon on the one hand by a few dominating personalities amongst the contractors, and on the other by changes in the current fortunes of existing railways. These fluctuations in railway-building were aggravated by the time which it took to promote and build a railway. Several years might elapse between the initial flotation on the Stock Exchange and the completion of the first stretch of line. This stretch generally involved the sinking of a large indivisible amount of capital which could not be cut without abandonment of the scheme as a whole. Once a start had been made on any important project, such as a trans-continental railway in Canada or the U.S.A., or a trans-alpine railway between France and Italy, the project had to be pushed on with even when it proved much more costly or less remunerative than had been expected. More and more capital was sunk in order to get some return on what had already been spent, with the natural result that investors were doubly shy of similar large-scale projects for some time to come.

This lumpiness of some types of investment is a third factor making for instability. The original decision to build, once taken, is irreversible, and investors are committed to the completion or abandonment of the project. Since they cannot readily cut their losses and cannot be certain to what scale of investment they are committed, they require all the more exuberance to be willing to finance the scheme. In practice, this means that it is easiest to obtain financial support near the top of the boom, when costs are rising and a break in prices is not far distant. The commencement of large schemes of capital expenditure at such times intensifies the boom; and the poor return which they earn in the ensuing slump keeps investment low and impedes recovery.

A fourth cause of instability lies in the fact that the demand for capital goods derives from the demand for consumable goods. Suppose that sales of motor cars amount to one million per annum and that 10 per cent. of the plant is renewed every year. Then if sales go up by 20 per cent. to 1·2 million, and if the extra plant to make additional cars has to be installed within the year, the total amount of plant to be built will be thrice as much as in previous years; instead of the normal 10 per cent. for renewals, it will be necessary to provide an extra 20 per cent. to meet the expansion in demand. The fluctuation in employment will obviously be much larger in the firms making machinery for motor car manufacture than in the motor car industry. The increase in the first will be 200 per cent.; in the second, only 20 per cent.

This last point may be put in a slightly different way by saying that the demand for consumption goods is at bottom a replacement demand, since a stream of these goods is constantly moving towards points of sale to replace the goods which are being consumed. The demand for capital goods, however, is at bottom an extension demand, designed to provide further supplies of consumable goods at some future date. The demand for consumable goods is largely governed by

custom and habit which exercise a stabilising influence; but the demand for capital goods is governed to a large extent by *changes* in the demand for consumable goods, and these changes are magnified out of all proportion in their repercussions on investment.

The problem of stabilising employment, to sum up, is fundamentally one of stabilising investment. If employment in the capital goods industries is steady, then total employment will be likely to be steady also. Unemployment is generally the result of a deficiency of investment.

## The Keynesian Theory

All this was well known before Lord Keynes published his "General Theory of Employment, Interest and Money." The peculiar contribution of that work was to set out more clearly the connection between savings, investment and income, and the part played by money and the rate of interest.

**Savings and Investment.**

Lord Keynes begins by defining savings and investment so that they are necessarily equal to one another. "Savings" are taken to be the difference between income and expenditure on consumable goods; "investment" is the value of the current output of capital goods together with the value of any addition to work in progress or to the stock of finished goods. On these definitions it follows that:—

(1) Savings=Income *minus* Expenditure on Consumable Goods.

(2) Investment=Output *minus* Output of Consumable Goods.

But since total income must be equal to the value of total output, and expenditure on consumable goods must be equal to the value of consumable goods produced (excluding additions to stock), the remaining items in the equations must also be equal. By definition, therefore, savings and investment must be equal.

This seems natural enough; one would expect equality between the value of the resources released by saving and the value of the resources absorbed by works of construction. But on other equally valid definitions (such as Lord Keynes used in his "Treatise on Money") savings and investment need not be equal. The definition of savings, which seems so straightforward, is in fact rather paradoxical It implies that if I receive my wages at 1 p.m. on Saturday, meaning to spend them at 2 p.m., they are part of my savings between 1 p.m. and 2 p.m. If I use them to buy a piece of furniture, they become part of the savings of the furniture salesman at 2 p.m., and when he settles with his wholesaler, carry on as part of the wholesaler's savings. In other words, any cash in our possession forms part of our savings. One corollary of this is that if more cash comes into existence, either through an expansion in the note issue or the creation of additional bank credit, the act of putting this cash into the hands of the public automatically increases savings. But such an act may leave unaffected the public's *intentions* about savings; those who receive the extra cash may have made up their minds to spend it at once. In one sense,

therefore—although not in Lord Keynes' sense—savings will be no higher than before.

However this may be, paradox is not a sufficient reason for quarrelling with definitions, provided the definitions are self-consistent and allow useful conclusions to be drawn. We can only draw these conclusions when we advance from the arithmetical to the causal. *Why* are savings and investment equal?

They are not equal because the rate of interest operates to make them equal. One might be tempted to think in terms of supply and demand and look upon the rate of interest as a price which rises and falls in response to changes in savings and investment. But the analogy is deceptive. If the economic system was so self-adjusting that all savings were automatically absorbed into investment at an appropriate rate of interest, there could be no lack of balance between the two. A reduction in investment, with habits of thrift unchanged, would be prevented by a fall in interest rates which would uncover new outlets for savings, and there would be none of those violent disturbances in employment which only too plainly do occur.

But if not the rate of interest, what *does* preserve equality between savings and investment? How does it come about that although habits of thrift are stable and investment is unstable in the extreme, total savings move in the same direction and with the same violence as total investment?

Lord Keynes' answer is, in brief, that it is through fluctuations in the national income, and corresponding fluctuations in employment and unemployment, that savings are kept in step with investment. An increase in investment creates additional employment and additional income, and out of this additional income come the savings required to finance the original act of investment. Investment does not wait on savings; it is savings that wait on investment. The more stable our habits of thrift, the greater the severity of the fluctuations in income and employment, since the greater is the change in income necessary to procure a given change in savings. As for the rate of interest, it is not affected by thrift; it is governed by the state of liquidity preference. And its rôle is limited to the influence which it exerts on the scale of investment.

Let us take the three variables in turn—investment, savings and the rate of interest—and examine the rôle of each of them.

(i) **Investment.**—It is not hard to establish that the level of the national income is closely associated with the scale of investment. Consider, for example, the following estimates for the United States in the years 1929 to 1941[1]:—

[1] The estimates are taken from Tables I and III of Mr. Stone's "National Income, Output and Expenditure of U.S.A., 1929-41," *Economic Journal*, June-Sept. 1942. The reader will find it instructive to plot investment against income for these years and sketch the curve along which they lie. This curve, which proves to be a straight line, represents the functional connection between investment (or savings) and income.

| | Investment $000m. | National Income $000m. |
|---|---|---|
| 1929 | 11·3 | 87·6 |
| 1930 | 6·4 | 76·0 |
| 1931 | 3·0 | 59·6 |
| 1932 | —2·2 | 42·8 |
| 1933 | —2·2 | 41·9 |
| 1934 | 0·8 | 50·1 |
| 1935 | 2·8 | 57·0 |
| 1936 | 5·7 | 66·9 |
| 1937 | 6·0 | 72·6 |
| 1938 | 4·3 | 66·1 |
| 1939 | 6·9 | 72·0 |
| 1940 | 9·2 | 79·8 |
| 1941 | 13·7 | 99·1 |

These figures bear witness to the enormous fluctuations that took place in the United States over these years. In the three years 1929-32 the national income fell by some 50 per cent.; and in the four years 1933-7 it expanded again by some 75 per cent. The fluctuations in investment—or, what comes to the same thing, in savings—were even more prodigious. In the two prosperous years 1929 and 1941, investment reached an average of $12,500,000,000; that is, the United States saved in those two years the equivalent of almost the whole of our pre-war National Debt. In the two leanest years, 1932 and 1933, investment was a negative quantity and the United States was actually drawing on capital.

Closer inspection of the figures shows a close parallel between the fluctuations in income and in investment. Not only do the two things move in the same direction and at the same time; the *magnitudes* of change bear a fairly constant ratio to one another. Roughly speaking an increase in investment by $1,000 m. would seem to accompany an increase in income by $3,500 m. The fluctuations in investment are, in fact, multiplied by 3·5 before they emerge as fluctuations in income.

Why do we get this "multiplier effect"? What is the mechanism by which a change in investment is translated into a change in income? If investment increases—for example, because of a great re-housing campaign—it will be necessary to employ more workers in the building industry and in all the ancillary industries—the timber and furniture trades, brick-making, the iron and steel industries and so on—which provide materials and fittings for houses. Earnings in these industries will increase by the amount of the additional investment. As earnings increase, so will expenditure. The spendings of those who benefit directly will create secondary employment in other industries—the clothing trades, for example—and these industries in turn will have more to spend. The increase in income to which the original invest-

ment gave rise, will be multiplied as prosperity spreads. But not indefinitely. The process of expansion will continue until a new equilibrium is reached with incomes so much higher than before that the increased scale of investment is matched by a correspondingly increased rate of saving. The investment will be financed out of the savings which it creates.[1]

This makes it a little easier to see why the " multiplier " can be given a fairly determinate value for any one country. Since savings and investment are equal, the " multiplier," which measures the rate at which income changes as investment changes, measures also the rate at which savings change as income changes. If the " multiplier " is 3·5, this means that out of an increase in income of $3,500 an extra $1,000 will be saved. Or, to put things the other way round, if income falls by $3,500, expenditure will be reduced by only $2,500. The size of the " multiplier " will obviously vary from country to country since savings do not respond in the same way everywhere to variations in income.

**(ii) Savings.**—To savings Lord Keynes allows a purely passive role. He argues that although it is open to each of us to decide how much of our income we wish to save, our joint decisions do not determine how much in the aggregate the community saves, any more than our joint decisions as to how much cash we want to keep with the joint stock banks governs total bank deposits.[2] So long as investment remains constant, savings must remain constant. All that the public can do is *spend* more or less ; alter what Lord Keynes calls their " propensity to consume " or (looked at the other way round) their state of thriftiness. This is not the same as *saving* more or less. For the act of spending affects other people's incomes and *their* savings. If a group of people set out to increase their savings they must begin by spending less on goods and services ; some trades will lose customers, and profits and employment will suffer ; those who lose money in this way will save less ; and this reduction in savings will counterbalance the original increase unless some fortuitous stimulus to investment arrests the contraction in employment and incomes. Greater thriftiness, if not accompanied by greater investment, merely impoverishes the community ; the increase in thrift is offset by a reduction in income, and the resources which are released by the thrifty are left idle for lack of investment. All that thrift can accomplish is negative if there is no machinery by which acts of thrift are translated into acts of investment.

[1] Savings and investment will be equal at each stage. But this equality will be achieved in the early stages (*a*) because consumers have not had time to spend their increased earnings and are holding more cash than they intend to ; and (*b*) because there will be some *dis*investment in stocks of consumable goods which manufacturers will wish to replace once they have taken on more workers. Full equilibrium is not reached until consumers have adjusted their habits of spending and manufacturers have adjusted their output.

[2] See above, page 328.

**Investment and Savings with Full Employment.**

To those who think of savings as an indispensable preliminary to investment (as it is for each of us), Lord Keynes may appear to put the cart before the horse. Before we, as a community, can save, we must invest; and if we do invest, this of itself will create the necessary savings. These seem paradoxical conclusions. Do they apply in all circumstances? Are we to deduce that we can increase the national income indefinitely by the simple device of steadily expanding investment?

The answer is that, in *money* terms, we can; but in *real* terms, obviously not. It is only in a situation in which there are still unemployed resources to be brought into use that further investment will increase *real* income. Once investment passes the rate which would sustain full employment, inflation results. The extra investment is financed by credit creation; workers who would otherwise find employment in making consumable goods are attracted into the constructional trades by the offer of higher money wages and are paid from advances made by the banks. The new money created in this way is saved in the sense that at any moment in time somebody owns it and has not yet spent it; but it is not saved in the sense that its owners intend to invest it rather than part with it in exchange for consumable goods. Total money incomes are inflated by the extra bank credits; and total savings are equally inflated, but only momentarily. Sooner or later the additional purchasing power will be spent. But the output of consumable goods will have reached the limit of expansion and will be threatening to contract because of the diversion of labour to the production of capital goods. The output of consumable goods, therefore, will fall short of the quantity which the public wishes to buy at current prices with the purchasing power at their command. In the scramble for these goods, prices will be forced upwards; wages will follow; it will be necessary, in order to maintain the scale of real investment, to have recourse to the banks for additional credits out of which to meet the increased cost of capital goods; and these additional credits will give a fresh impetus to inflation.

The situation, in short, becomes one in which producers of capital goods and consumable goods are in active competition with one another for command over real resources, and their joint demand is greater than can be met from the resources available. The result of their tug-of-war is to bid up the prices both of the factors of production and of finished goods, the makers of capital goods sustaining their demand through the creation of additional bank credits, and the makers of consumable goods sustaining their demand because these credits ultimately swell their takings.

It is when we approach a situation of this kind that savings begin to fulfil their true function of releasing resources for employment in works of capital construction. *Then* investment can increase without provoking inflation, only if the demand for consumable goods is held in check by fresh saving. At this stage, too, the rate of interest does begin to operate to keep savings and investment in step. For if the

banks, taking fright at the prospect of inflation, refuse to create additional credit, it will be by raising interest rates that they seek to ration credit amongst competing claimants. Each successful project of investment that elbows its way forward to receive the financial backing of the banks or of private investors will simultaneously oust some other project which does not offer a large enough prospective return to cover the interest payable. Investment can be held constant in spite of a rise in the marginal productivity of capital if the rate of interest is allowed to rise simultaneously.

**(iii) The Rate of Interest.**—Interest rates, according to Lord Keynes, are the price not of savings but of liquidity; the complex of interest rates is governed by the state of liquidity-preference.[1] Interest rates will be lower, therefore, the larger the supply and the smaller the demand for money; and those rates of interest will be most responsive to changes in the supply of money which are paid on short-dated, highly liquid securities.

The supply of money is within the control of the banking system; by a suitable credit policy, therefore, pressure can be brought to bear to increase or reduce interest rates. This pressure is not likely, however, to be effective of itself in bringing about a parallel change in long-term rates of interest. Investors must first be convinced that cheap money has come to stay.[2] If they believe that the fall in short-term rates is "unnatural" and due to an "unsound" policy which will later be reversed, their reaction may be perverse: they may give expression to an increased preference for cash by offering lower prices for bonds and forcing up long-term interest rates. Long-term rates reflect long-term views. They are influenced more by propaganda than by open-market operations. Moreover, they are less vulnerable to pressure in a downwards than in an upwards direction. The insensitivity of long-term rates cannot be overcome by reducing short-term rates indefinitely (certainly not below zero); but it can be overcome by an indefinite increase in short-term rates. Even more important in setting a lower limit to long-term rates is the part played by uncertainty about the future of interest rates, and the consequent danger of capital depreciation. The lower the rate of interest, the higher is the price of bonds, and the smaller is the annual yield in relation to the risk of capital depreciation through an unforeseen recovery in interest rates. If the entire weight of the banking system is behind efforts to reduce interest rates in a slump, they may still fail. And historically, the banking system has often sought to restrict credit at such times and to expand it in times of boom.

The theory of liquidity preference has been the subject of much controversy amongst economists. Lord Keynes' original version made it appear as if interest-rates were affected neither by savings nor by investment. But the rate of interest is in fact affected

[1] See above, pages 258, 343 *et seq.*

[2] See above, page 259.

by both: by savings because, with full employment, the rate of interest necessary to prevent inflation will be high or low according as the propensity to consume is high or low; by investment, because the demand for money is at all times closely associated with the funds that must be held before projects of capital construction are embarked upon.

**Investment in a Boom.**

Even if long-term rates of interest were completely under control, this of itself would not ensure stability of employment. The quality as well as the volume of investment changes in the transition from slump to boom and projects are financed which, looked at with due prudence, cannot hope to give a return equal to the current rate of interest. If the rate of interest is increased to meet a multiplication of such projects, the effect may be only to assist them in supplanting other more worth-while projects. In due course when the misdirection of investment becomes apparent, it may be too late to check the set-back in investment by anything short of a nil or negative rate of interest; for the projects which were shelved will take some time to revive, and the immediate effect of the crash in speculative investment will be to create an atmosphere most unfavourable to further investment, however attractive in normal circumstances.

**A Corollory of the Keynesian Theory.**

It follows from the Keynesian theory that the difficulty of maintaining full employment tends to increase with the standard of living of the community. If the standard of living is low, a given reduction in income tends to have less effect on the level of consumption because there is less scope for further stinting. A reduction in investment, therefore, is more readily counterbalanced by a reduction in savings. But if there is a high standard of living, it is easier to make economies in expenditure and maintain savings, and unemployment tends to spread over a wider area. Of much more importance, however, is the fact that the scale of investment in relation to total income and employment is much higher. Capital accumulation, therefore, plays a much larger part in the life of the community, and employment is consequently much more at the mercy of oscillations in investment.

**The Future of Investment.**

The trade cycle arises because of a periodic deficiency of investment. But there may also be a chronic deficiency of investment; there is a danger that the rate of interest may be too high to allow us to take full advantage of the community's thrift, and that the riches which dispose us to save prove our undoing in unemployment.

How great is this danger? We know that it is hard to push interest rates below 2½ per cent.; we know, too, that there was persistent unemployment in the ten years before the war on an unprecedented scale. These facts are consistent with chronic under-investment. But are there not almost unlimited openings for investment in a well-

managed community? And if not, why should it be impossible to dissuade people from saving so much?

Let us try to see the problem in perspective. The meagre savings of 1842 would have met the cost of the war for a week in 1942. Yet out of those savings came the apparatus of modern industrial life —the railways, the steamships, the gasworks, the mansions, terraces and cottages of Victorian England. In the nineteenth century there were new lands to be opened up, new industries to be built, new families to be housed. In the undeveloped countries the scope for capital outlay was enormous. There were recurrent depressions when satiety seemed to have been reached. But these were rather intervals for the recovery of fresh appetite after some peculiarly sick-making misdirection of investment. They did not result in extensive unemployment because they were generally accompanied by an interruption in the surge of population into the new countries, the fall in immigration compensating in part for the fall in investment; they did not even cause serious unemployment in the old countries of Europe, because the exodus was mainly from the rural areas, and a temporary decline in emigration meant land hunger and under-employment in the villages rather than unemployment in the towns.

In this frontier investment, capital had an abnormally high marginal productivity. This was so, partly because many of the chief inventions of the industrial revolution were largely associated with transport and communications, and partly because these inventions involved the extensive use of capital. The reduction in transport costs meant that areas of the earth which had previously been relatively inaccessible could now be used to furnish Europe with foodstuffs and raw materials. Capital was needed to produce both foodstuffs and raw materials. But it was also needed for the railways and ships that carried them. In Europe itself the improvement in communications also called for an enormous capital outlay, chiefly on railways. Then there was capital to be found for manufacturing plant for all the new industries and the new techniques.

Throughout the nineteenth century investment was sustained by an expanding market based upon an increasing population. The yield from a given asset did not necessarily decline with obsolescence but might be higher twenty years after construction than five years after. Houses, for example, were an irreproachable form of investment. With the semi-automatic increase in the demand for houses that went with the growth of population and its concentration in cities, no landlord had reason to fear that his property would long stand empty; he might, with luck, draw a higher rent as time went by and its nearness to the city centre increased its attractiveness.

Nowadays, things have changed. There are no frontiers still to be explored. In India, China and Africa there may be scope for investment on an enormous scale; but it would not yield the fabulous returns earned in the "empty" countries of a century ago—America, Canada, Australia, etc. The new industries are less greedy of capital, in relation to modern standards, than those of Victorian days. No

form of *private* investment in the twentieth century parallels the railway building of the nineteenth—although *public* investment in roads, dams, irrigation schemes, and so on, does. The rate of population growth has slackened perceptibly, and investment has, therefore, come to depend more on the rate of obsolescence of plant and property and less on the semi-automatic growth of the market. Above all, the investment potential of twentieth century nations is on a tremendous scale. The United States can vote a war budget of £20,000 million, which is one hundred times as much as Great Britain was able to lend or invest abroad at the peak of her foreign investment. It is easy to employ such enormous sums in war; is it so easy in peace?

Provided the State is willing to do "unorthodox" things, there is no reason why it should not be. If a country's savings are too great, they can always be given away. They can be given away by voting loans, free of interest, to backward countries; or if necessary, by free grants to such countries of surplus food, or steel, or machinery. There is plenty to stir the imagination in the idea of endowing Europe with a good road system, or Africa with hydro-electric power, or China with levées to prevent floods, or India with schools. The obstacles are psychological: reluctance to offer charity elsewhere than at home, so long as unemployment and long hours and poverty are still the lot of millions in the rich industrial countries. But if these obstacles are insurmountable—and they need not be—the State can equally well distribute purchasing power to its own people. It can either give money, or paid holidays, or food and clothes to selected categories of workers. It is true that this also will create psychological difficulties amongst the groups who do not qualify for assistance. If these difficulties in turn are insurmountable, the State can fall back on enriching its cities and villages with finer buildings, on slum clearance schemes, or the complete re-planning of towns.

All of these things would add to the National Debt. If the State tried to finance all of them by extra taxation, it would be removing purchasing power with one hand and releasing it again with the other. But there is no reason why we should fear an increasing public debt so long as the National Income is increasing pro rata; and merely by preventing unemployment, the State will be helping to increase the average level of the national income and providing itself, out of the additional revenue which this will bring, with part of the cost of its "unorthodoxy."

The alternatives are two-fold: to conduct an anti-savings campaign; and to rely on private investment. The first is not likely to be very successful. For people do not save out of sheer miserliness; they have generally solid reasons grounded on self-interest. For one thing, savings are a ladder—a method of rising to a higher social class; this was perhaps why the Victorians, who respected success, commended thrift. So long as social mobility depends upon saving, habits of thrift are not likely to yield to propaganda.

As for private investment, this is of limited force. Much of it is in housing. Here, in spite of the diminishing pressure of population growth, the marginal productivity of capital has been sustained by obsolescence, because the houses in which people once chose to stay are considered inadequate by modern standards. Much private investment also is in public utilities, and again invention has widened the scope for investment: for example, by providing new uses for electricity. But an ever larger proportion of investment is undertaken by the State or by bodies prompted or controlled by the State. This is true of roads, railways, air transport, public utilities, housing, and even factory-building. Thus it is by administrative decision rather than by reference exclusively to interest rates and prospective yields that investment is coming to be determined. This of itself should help to stabilise investment. But it also puts on the State an added responsibility for seeing to it that the aggregate volume of investment is sufficient in relation to the potential savings of the community.

# PART VII—THE STATE AND THE ECONOMIC ORDER

## CHAPTER 29

## THE STATE AND INDUSTRY

A CENTURY ago it was almost an axiom that interference by the State in the conduct of industry was mischievous and inadvisable. The business of government was to preserve law and order, to defend the country from invasion, and, when opportunity arose, to extend the boundaries of empire. The business of industry was to make things cheaply. Interference by industry in politics was fatal to good government, and interference by the State in industry was fatal to efficient production. Business influences, it was argued, are apt to corrupt politics; and political influences are apt to corrupt business. The wisest course for a govenment to follow was to pursue a policy of laissez-faire; to abstain from a blundering intervention and leave industry to its own devices.

The movement in favour of laissez-faire drew strength from two more or less distinct groups. There were those who disliked government interference on principle, because it limited the freedom of action of the individual; and there were those who disliked such interference as they encountered in practice, because they felt it to be foolish or unjust. The first group appealed to the natural rights of the individual; the second, to their business experience. Both groups —philosophers and business men—were united in opposition to authoritarianism. They challenged "the right divine of states to govern wrong" because of a common dislike of bureaucratic restrictions on freedom of opinion and of enterprise.[1] Not that these early "liberals," having taken to the path of individualism, had any clear idea where it led. The business men fought hardest—generally from motives of self-interest—against particular restrictions which they found irksome or costly; but they were often disposed to put other restrictions—no less intolerant and monopolistic—in their place. The philosophers fought against encroachments by the State on individual liberty; but they defended rights of property which could be as real a menace to the liberty of the unpropertied as any govern-

[1] Thus it was no accident that religious toleration and laissez-faire grew up side by side. For both were rooted in the experience and outlook of a rising social class, concerned to get on and impatient of authority. In economics, this class demanded free trade in goods and the suppression of privileged monopolies; in religion, it demanded free trade in opinions and the suppression of religious discrimination. Competition in industry would eradicate inefficiency—or at least increase profit; competition in ideas would eradicate untruth—or at least increase the force of propaganda against the governing class.

mental despotism. Moreover, although both groups were united in condemning State interference as *undesirable*, it was not until laissez-faire found a third group of supporters—the economists—that interference came to be thought of as *unnecessary*. It was the economists who, by making the virtues of laissez-faire intelligible in a synthesis of principle and experience, paved the way for the final triumph of the policy.

**The Philosophers.**

The arguments of the philosophers took one form in France and another in Britain. In eighteenth century France, a school nicknamed the " Physiocrats " exalted " the laws of nature " as the touchstone by which society should be organised and compared these natural laws with the " conventional " laws of states to the disadvantage of the latter. They sighed for some kind of return to nature and were profoundly distrustful of the achievements of organised society. Personal liberty was hamstrung by the threat of the Bastille; commerce was crippled by vexatious restrictions on the movement of goods from one part of France to another; and the immense natural wealth of the land was insufficiently exploited because of a wasteful tax-system and an outworn feudalism. Only by the removal of unnecessary restrictions and the immediate adoption of a policy of laissez-faire could France be saved from a threatened catastrophe.

In Britain, the philosophers defended their individualism by an appeal to the "natural rights" of man; the right, for example, to " life, liberty, and the pursuit of happiness." Man, they argued, is an individual first and a member of civil society afterwards; and he can be called upon to surrender his rights as an individual only when these rights are a menace to the equal rights of his neighbours. Each citizen barters some liberty of action for the common protection to life and property afforded by the law of the State. But he gives up no more of his natural liberty than is necessary to secure that protection, and undertakes to obey the law only so long as it enforces his natural (individual) rights. He is under no moral obligation to submit to arbitrary acts of tyranny or to laws conceived in the interests of privileged groups. If the State exceeds the powers which he and his fellow citizens have tacitly delegated to it, he is entitled to resist its authority by appealing to the higher authority of natural right. It is the State's business to protect, not to usurp, the natural rights of its citizens.

In practice, the rights which were given most emphasis in this theory of the State were those which found most favour with the rising middle class—rights of property. Everyone, it was claimed, should have the right to enjoy the full product of his labour and to do as he wished with his own. The less the State interfered with property, therefore, the better. In imposing excessive or arbitrary taxation, in placing burdensome and unnecessary restrictions on industry, in making grants of monopoly to privileged traders, in seeking to fix

prices, wages, or rates of interest, the State was overstepping the narrow limits of its proper functions and infringing natural rights. The individual—and particularly the individual property owner—was justified in challenging any such infringement.

**The Business Men.**

Business men took their stand on the lower ground of self-interest. They might protest their regard for the public good, but the grinding of axes could generally be heard in the background. They were not averse to government regulation on principle, and indeed, until the Industrial Revolution, had favoured rather than opposed such regulation. So long as the market which they supplied was no larger than a mediæval town, and so long as each town was forced to defend its own interests (or even its independence) against the uncertainties of the outside world, the merchants were forced to band themselves together in " guilds " and submit to collective regulation of their trade. Such regulation was accepted freely, since it was conceived in the common interests of the guild and administered by its members. Similarly in foreign trade, so long as opportunities of profit were restricted by the narrowness of the market, and so long as the risks to life and property made State protection more or less indispensable, State regulation was taken for granted. And in industry, whenever large amounts of capital had to be sunk in times when capital was scarce, for a very uncertain return, the need for State protection and assistance was generally recognised. But as the market expanded and capital accumulated, the attitude of the business community began to change. The growing security which surrounded trade reduced the chief need for joint action—that of mutual defence. The accumulation of capital increased the number and the ambitions of traders, making them more venturesome and more impatient of restrictions. Above all, the expansion of the market stimulated their acquisitiveness by providing them with new and attractive opportunities of profit, put a premium on their initiative by substituting varying for customary trade connections, and enormously reinforced the difficulties of regulation by making business a much more complex and a much less personal affair. The market acquired a growing ascendancy as the mechanism by which consumers and producers were brought together. And with the rise of the market went also the rise of competition and of individualism. The traditional ways of co-operation between traders were sapped as they came to depend less and less on their neighbours and more and more on the prices at which they bought and sold. The more they thought in terms of the market, the more they were tempted to ignore other social bonds and obligations for the sake of buying cheap and selling dear. They fought for their own hand and resented interference by outside bodies, governmental or otherwise.

To some extent, their resentment was simply the outcome of their individualism; they did not like to be reminded of their obligations to society. But it would be foolish to represent the governments

of the seventeenth and eighteenth centuries as the champions of wide social interests against the selfishness of business men. The truth is that there was no government machinery adequate to the tasks of State regulation. There was no trained civil service and no disposition to create one. The business man could argue—and with force—that it was to his initiative and enterprise, and to his first-hand knowledge, to which no mere official could have access, of the methods and processes of industry, that society must look for the successful conduct of industry. He could suggest—with even more force—that all, or nearly all, government intervention in the struggle for profit was dictated by petty jobbery, sectional interests, prodigality or ignorance of political economy; and that, even when the need for intervention was admitted, governments could not be relied upon to act promptly and effectively.

**The Economists.**

The contribution of economists to the controversy developed as a critique of "Mercantilism," the system of State regulation which preceded that of laissez-faire. The views of Mercantilist writers were rarely expressed coherently and systematically and were even more rarely implemented by a consistent economic policy. But they were always in the background as the mental setting against which policy took shape.

**Mercantilism.**

The core of Mercantilist doctrine was the prime importance to a nation of procuring a favourable balance of trade. By selling more to foreign countries than it purchased from them, a country would be able to demand payment of the balance in gold or silver bullion, and so add to its stock of the precious metals. Such an addition was considered desirable for a variety of reasons.

First, it increased the country's financial strength. It formed a war-chest which could be drawn upon, much as our foreign investments can nowadays, to meet the exceptional cost of fitting-out an army and maintaining it abroad; and in times when there was no funded National Debt and credit facilities were comparatively primitive, such a war-chest of ready money was altogether indispensable. It is true that the precious metals procured by trade did not flow directly into the government's exchequer. But they lay easily to hand in private hoards, just as our foreign investments do in peace time, and could be requisitioned by special taxes and forced loans. It is true also that some Mercantilists were so impressed by the importance of the precious metals that they tended to ignore other kinds of wealth. But this was only because of their obsession with war potential; they did not exclude land and buildings from the nation's wealth, but only ranked them below the precious metals in military value.

The undesirability of an unfavourable balance was also stressed. A persistent drain of the precious metals was regarded with much the

same disfavour as we feel towards its modern equivalent—a flight of capital; and laws were passed to prevent "good English money" from leaving the country, much as modern countries institute exchange restrictions so as to keep their money "at home." The Mercantilists realised, too—what in modern times we have been reminded of by hard experience—that an unfavourable balance tends to bring with it trade depression and unemployment. In the Mercantilist period (the sixteenth to eighteenth centuries), price-relationships were dominated by custom, and wage and other costs were consequently rigid, so that adjustment to an unfavourable balance was particularly difficult. An active balance had the merit of giving a mildly inflationary impetus to industry, easing any downward pressure on wages, and bringing about an expansion in profits and trade activity. Not that the Mercantilists put their point of view in quite so sophisticated a way; many of them, indeed, had in mind nothing more subtle than a crude, self-interested protectionism. But the many schemes put forward for "setting the poor on work" by excluding imports or by depreciating the rate of exchange, suggest that the Mercantilists were genuinely anxious about unemployment and that they connected it vaguely with the trade balance.

With a view to procuring a favourable balance of trade, the Mercantilists advocated the close and constant supervision of the economic life of the nation. They wished duties to be imposed on imports and bounties to be granted on exports. They wished laws to be passed against the consumption of foreign luxuries, which took money out of the country and tempted men from a patriotic and puritanical love of toil. They wished wage rates to be low so as to make exports cheap and so as to provide the worker with an incentive to exert himself. They wished all productive resources to be utilised to the full; no land to be left uncultivated and no labour unemployed. In short, they wished for the maximum of production and the minimum of consumption, the difference to be accumulated in bullion. If the precious metals could be obtained by less roundabout methods, so much the better. A country might have gold mines within its boundaries; if so, the mines should be exploited to the uttermost. It might win colonies which were rich in the precious metals; if so, the colonies should be stripped of their riches. If the colonies were unfortunately barren of the precious metals, their natural wealth might still be turned to account by bringing home colonial produce for manufacture or for re-export at a profit. Colonies were sought, not for settlement, but for exploitation. They were to be developed, not for their own sake, but as estates of empire supplying the mother country with raw materials which it lacked and with markets for the manufactures which it wished to sell.

The values of the Mercantilists were primarily military. They put guns before butter, defence before opulence. It was their insistence on military strength which led them to advocate a large population, whatever its standard of living. For the same reason, they discouraged trade with foreign countries in foodstuffs and raw materials.

To depend upon imported grain was to put ourselves at the mercy of the countries which fed us; to export wool was to allow other countries to compete with us in the manufacture of woollen goods.[1] The export of manufactures, however, was an almost unqualified advantage; these exports were thought to create employment in the working up of native materials, to confer no advantage on foreign countries (either by giving them a dangerous control over supplies, or by enabling them to compete in other markets), and to bring in payment a store of bullion which could serve as the sinews of war.

**The Downfall of Mercantilism.**

The corner-stone of Mercantilism was the doctrine of the favourable balance of trade. Once refute this doctrine and the whole structure begins to crumble. But refutation is easy; we need only invoke the Quantity Theory of Money. For if a favourable balance, by adding to the money-supply, merely causes prices to rise, and if the rise in prices automatically puts an end to the favourable balance, then the State is pursuing a will o' the wisp in trying to procure a continuously favourable balance of trade. Once it is seen that gold flows are automatically reversed by the inflation or deflation which they set in motion, the vigilance of the State over the trade balance appears ineffectual and superfluous.

State regulation of industry seemed even less necessary after Adam Smith had laid bare the workings of competition. For it was not only the movements of gold which were governed by a self-regulating mechanism; the entire economic system, if left to itself, would function automatically in the best interests of the community. There was no need, for example, to fix fair prices or wages; free competition would drive both to their natural and proper level. There was no need to offer a subsidy to one industry or to cripple another by taxation; free competition would put both on a fair and equal footing. There was no need to shackle the natural cupidity of business men—indeed it was almost impious; for the "invisible hand" of market forces would transform cupidity into an unintentional benevolence. It was not the charity of bakers which assured us of our daily bread, but a much more calculable and powerful motive—their self-interest. It was this motive which made the economic world go round without the meddlesome intervention of governments. Self-interest, spurned by the moralists as a vice, was exalted by economists as the driving

[1] Compare, for example, the following passage from Heywood's "King Edward the Fourth":—

*Rufford:* Mistresse, I fear you have forgot my suit.
*Jane Shore:* Oh, tis for a license to transport corne
From this land, and lead, to foreigne realmes.
I had your bill; but I have torne your bill;
And twere no shame, I think to teare your eares,
That care not how you wound the commonwealth.
The poor must starve for foode, to fill your purse,
And the enemy bandy bullets of our leade!

force behind material progress and an adequate substitute for government regulation.

There was an element of idealism, too, in the arguments of the economists. They could urge the value of foreign trade as a force making for peace. Trade was of advantage to both parties, not, as the Mercantilists believed, to one only; it cemented international friendships while the beggar-my-neighbour policies of Mercantilism fostered international jealousies. By releasing its trade from the restrictions of a "mean and malignant system," therefore, a country could make a great contribution to international goodwill. On the other hand, by maintaining the restrictions, it invited such disasters as the loss of the American colonies in 1783. In domestic politics, it was possible to defend free and unrestricted competition on an equally high plane of idealism. Competition represented a technique for settling conflicts of economic interest which compared favourably with lobbying by vested interests in a corrupt Parliament. In comparison with the granting of monopoly powers and privileges by governments, it was a step towards equality of opportunity. For the blunders, the routine, the red tape, the indecision and timidity of a bureaucratic machine, it substituted the initiative of a class of entrepreneurs, spurred to quick decision by self-interest and by intimate knowledge of their trade. And by exalting individualism in business, it turned the energies of restless and self-seeking men into the comparatively harmless channels of commerce and colonisation.[1]

The downfall of Mercantilism after Adam Smith is not to be explained solely in terms of a change in ideas, but also in terms of a change in circumstances. Defence motives were giving way to money motives. In the comparative security of eighteenth century Britain, a man's social standing was no longer measured in terms of his value as a warrior, but rather by his possessions. Trade was increasingly attractive because success brought social recognition, and increasingly prosperous because a strong central government had given order and unity to the country. Men's interests turned more and more to the exploitation in trade and industry of the geographical discoveries made by explorers and of the technical progress made by scientists. The arena of power was shifting from politics and religion to business.

As the danger of war, foreign and domestic, receded it was less necessary and less expedient to build, in the words of Adam Smith, upon the "compulsion, restriction, and monopoly," of a defence economy. Less necessary, because considerations of national power

[1] Compare G. M. Young (*Sunday Times*, 14th January, 1940):—
"People whose chief object in life is wealth are as a rule much less intelligent than people whose object is power, but, on the other hand, are far less disposed to do active wrong to those beneath them or about them. They may, from negligence or selfishness or mere thick-wittedness, do much harm . . . but it is harm of a kind which is largely controllable by public opinion, by legislation, by science."

declined in importance.[1] Less expedient, because expanding trade of itself brought immense financial power to a nation. Dependence upon foreign trade had its dangers; but these dangers could be met out of the profits of trade at the cost of a few battleships or by subsidies to continental allies. Economic power, based on laissez-faire, could be more than a match for military strength without financial backing, especially if the tendency of laissez-faire was to stimulate the industries which are most useful in making war materials, and the skill and habits of mind which are most useful in handling them.

The fact is that the Mercantilists dug their own grave by their exaltation of foreign trade. Mun was preparing the way for Adam Smith when, in 1623, he called his book "England's Treasure *by Foreign Trade*." He and Adam Smith were agreed that "defence is greater than opulence," but they differed in the emphasis which they laid upon defence, and in the methods by which they hoped to promote it. They differed, partly because of Adam Smith's much superior powers of logic and observations; but also because the time was ripe for new principles and because people were ready to listen to them. Mun, for example, might have accepted the Quantity Theory of Money; but he would not have attached much importance to it so long as he regarded the precious metals as stores of value, rather than as media of exchange; it was the growth of banking, and of alternative investments, which made men less anxious to obtain supplies of the precious metals. Again, Mun and his contemporaries were preoccupied with an unemployment problem which had ceased to trouble the late eighteenth century; they could not contemplate deflation with the optimism of Hume and Adam Smith. A policy of laissez-faire, in short, met the needs of the moment, not necessarily the needs of all time or of all countries.[2]

**Criticism.**

The case for laissez-faire, in international as in domestic trade, was largely a homily on the merits of cheapness and the demerits of government control. It was assumed that no means to the economical satisfaction of consumers' wants would be overlooked by business men set on their private advantage. Free competition was an automatic device for the extermination of inefficiency; free trade, a guarantee that the utmost profit would be reaped from the generosity of foreigners in offering cheap imports and in providing a remunerative market for exports. It was unnecessary for the government to regulate and plan industry, because prices already balanced supply and demand, cost and utility; pricing was an adequate substitute for planning.[3] It was also unwise for the government to intervene;

[1] It is worth noticing that from 1815 to 1914 (the heyday of laissez-faire) Britain was engaged in only one major European war—the Crimean.

[2] There was a common saying amongst foreigners in the nineteenth century that laissez-faire was made, like other Manchester goods, for export.

[3] For a development of this line of thought, see pages 138-40, 143-4.

governments were incapable of improving upon the delicate machinery of the market, and were likely, by their intervention, to throw the machinery out of gear, or to impose the dead hand of routine, or to act as the puppets of "sinister interests." By devolving the tasks of enterprise upon the business community, they could enlist stronger motives and more intimate knowledge in the planning of industrial output. And that output, produced at the lowest possible cost, would conform to the free choice of consumers as expressed in their purchases, not to some ideal imposed by an authoritarian government.

The case for laissez-faire was a strong one; it was realistic in its appeal to experience, and idealistic in its appeal to individualism. Historically, too, laissez-faire was a highly successful, if unnecessarily harsh, policy; the material progress made in the nineteenth century was immense, and it would be foolish to belittle it. But the policy had its drawbacks, and the arguments on which it rested, their flaws.

(a) **Competition is not Perfect.**—We have repeatedly had occasion to underline the growth of monopolistic elements in modern industry.[1] The free competition assumed by the advocates of laissez-faire is steadily diminishing—if, indeed, it ever really existed. On the one hand, the imperfection of the market results in wasteful advertising, duplication of services and of plant, the insulation of inefficient firms from competitive pressure, and the excessive multiplication of brands and types of product. On the other hand, the increasing scale of enterprise leaves a few giant concerns to divide the market, with no guarantee to consumers either that the monopolists are really efficient or that the benefit of reductions in costs will ever be passed on. The State has thus a double motive for intervening. It can hope to organise industry more efficiently than private enterprise, and it can ensure that the spoils of monopoly are not left in private hands. The wastes of imperfect competition can be eliminated by extensive rationalisation—by shutting down redundant plant, standardising output, and concentrating production in the most efficient firms. The extortions of monopoly can be prevented by setting up public corporations, managed by trained business men and aiming, not at maximum profits, but at efficient public service. Such public corporations—the Central Electricity Board or the Port of London Authority, for example—have many of the merits generally attributed to private enterprise. They can draw, for example, on men who have proved their ability outside the Civil Service; and they can pursue a bold policy of capital expenditure without constant recourse to the Treasury or to Parliament for approval in detail. Indeed, they can lay claim to superior efficiency, for they are free from the distraction of fighting trade rivals, and can devote themselves exclusively to the interests of consumers without copying the practices of private monopolies in suppressing patents, persisting in the use of obsolete plant, and so on.

[1] See above, Chapters 9 and 14.

The growth of monopoly creates a strong prima facie case for government intervention. But such intervention need not necessarily take the form of government operation of industry either through a public corporation or a department of State. It may be wiser to restrict government action to the fixing of prices and terms of sale; or to the regulation of monopolistic practices by a Trade Commission armed with wide powers of investigation and supervision; or to the mobilisation of public opinion by giving the maximum publicity to any abuse of monopoly power. Each of these expedients has its difficulties,[1] but they may be preferable to outright nationalisation. Public corporations, for example, are not likely to improve efficiency except where the size of the optimum firm is large and where its market coincides with some convenient administrative unit. Where there are strong economic forces—the need for energetic and resourceful management, for instance—making against large-scale enterprise, or where, as commonly happens, the optimum unit lies midway in scope between the central government and a local authority, a public corporation is beset with difficulties from which private enterprise is likely to be free.[2] Prices may still be reduced, especially if the danger of monopolistic exploitation is serious, but the gain to consumers will accrue out of the excessive profits of the monopolist, not from the greater efficiency of the public corporation. To enjoy the best of both worlds, it might be easier to leave the monopolist, with his more flexible organisation, in control of production, and, at the same time, to fix maximum prices for his output.

Where a private monopoly is replaced by a department of State or by a municipality, the gain is still more dubious. Democratic governments and their officials are constantly under the necessity of explaining and defending their decisions. But sound practical judgments are not always easy to explain; they may be intuitive and incapable of exposition in a speech or on paper. A glib memorandum, therefore, tends to carry more weight than the unsupported wisdom of men of affairs.[3] At the same time, the necessity to offer a satisfactory defence of public policy inculcates an excessive caution. Large risks, however justifiable, must result in occasional failures which can be seized upon by political opponents to discredit the

---

[1] For a full discussion, see A. C. Pigou: "Economics of Welfare" (4th edition), pages 336, *et seq.*

[2] These difficulties are most obvious where technical factors require a constant re-adjustment in the area under the control of a single enterprise—as, for example, in the supply of electricity. (*cf.* Pigou, *op. cit.*, page 402).

[3] Compare Hawtrey: "The Economic Problem," page 339: "To express in language even the decision itself is an effort: to express the grounds on which it is taken would often be a formidable exercise in both psychology and literary composition. . . . There is a tendency for any official hierarchy to be limited to those decisions that can be readily communicated in language from one functionary to another." (Quoted by Pigou, *op. cit.*, page 388).

government. In addition, governments are more easily deterred from large-scale experiments by the need to consider the interests of the whole community, than are private entrepreneurs who, having less at stake, can put their entire fortune to hazard without jeopardising the wealth and credit of society at large. A democratic government, in short, is less responsive to the hope of gain, and more alarmed by the fear of loss than private enterprise; its officials, in consequence of their training, are less open to new ways of thinking and less given to the making or exploitation of new inventions and new methods of organisation.

Not that failure to take risks is, in all circumstances, to be deplored. Many risks taken by private enterprise—particularly those which are borne by uninformed investors—are foolish on any showing, and merely result in a waste of capital. Many risks arise out of the competitive system itself, and are due to each business man's ignorance of what other business men intend to do. When the State plans, it may be comparatively slow and cautious in reaching decisions; but just because its plans are centralised and comprehensive, taken in the light of fuller knowledge and with regard to wider interests, it has less need to take risks and can take them with more foresight than the individual entrepreneur. Many risks, which fructify in real mechanical progress, may bring slight gains in comfort to the many at the cost of a total loss of livelihood to the few. Can the taking of such risks be safely left to uncontrolled private enterprise? As for inventions, these are made increasingly in the research departments of large firms which, in size and organisation, are comparable with public corporations. The difficulties confronting a government in encouraging invention are thus diminishing, although the difficulties confronting governments in trying them out before their commercial value has been established are still real enough.

When a government ceases to be democratic or where, as in war-time, it has a single aim to which everything is subordinated, many of the above considerations lose much of their force. The achievements of dictatorships or of war-time governments cannot be readily compared with those either of private enterprise (competitive or monopolistic) or of democratic governments; for the sacrifices to which people will submit, or which they willingly accept, under a dictatorship or in war-time are immensely greater than those which a democracy would contemplate in peace-time. The task of economic organisation, for example, is enormously simplified by the sweeping aside of vested interests and the suppression of internal conflicts. A dictatorship can plunder whatever classes can be represented as enemies of the people, can conscript labour whenever it sees fit, and can, by suitable propaganda, represent as a resounding success what in a democracy would be denounced as a shocking muddle or a costly failure.

In any discussion of government operation of industry, therefore, we must be on our guard against arguments based upon experience in other countries or in war-time. It is sufficient to observe that the case in favour of nationalisation of industry, so far as it rests on purely

economic grounds, is strong—increasingly strong—but by no means overwhelming. Some industries are riper for nationalisation than others; and public control would often be preferable to public operation.

**(b) Laissez-faire creates Unemployment.**—Whether or not private enterprise is an efficient system for making things cheaply, it certainly offers no guarantee of full employment for the productive powers of the community. At the bottom of a slump one man in three may be unable to find employment, and even in times of "boom" unemployment may still exceed 10 per cent. of the population seeking work. If it can be shown that this unemployment is the direct outcome of a system of private enterprise, and that it cannot be dealt with effectively within the limits of such a system, then we need go no further in search of arguments against the system. No economic system can stand up to intense and prolonged unemployment and survive. For in the first place, unemployment is a senseless waste of productive power; it reduces the National Income below what it might otherwise be. But the evil of unemployment extends far beyond the loss of income which it occasions. It ruins a man's self-respect by depriving him of an honourable opportunity of earning a living; it creates a sense of frustration and finally of uselessness; it saps his powers of concentration and his capacity for normal enjoyments; it makes for tension within the family and within the community; and it leaves men apathetic to ordinary social activities and duties, or ready to lend a willing ear to violent expedients for regaining their status and their sense of purpose.[1]

We have already seen that an uncontrolled system of private enterprise necessarily gives rise to unemployment. But there are various ways by which this unemployment could be reduced without abandonment of the system. Let us take the various expedients which have been suggested, one by one.

**(i) A Public Works Policy.**—If fluctuations in employment are chiefly due to fluctuations in investment, the most obvious way of stabilising employment is by stabilising investment. This can be done by offsetting fluctuations in the capital outlay of private enterprise by corresponding, but inverse, fluctuations in the capital outlay of public bodies. By postponing works of capital construction from times of boom to times of slump, the State can make both

[1] An investigation made between 1937 and 1939 showed that roughly half of the unemployed workers in the age-group 18-24, in Glasgow, Liverpool, and Cardiff, had had less than six months' work in the past three (comparatively prosperous) years; and that most of these workers took no interest in politics beyond speculating on the chances that a European war would bring them a job. Unemployed middle class workers are less apathetic: they played a leading part, for example, in Hitler's rise to power in 1931-33, when the unemployment rate in Germany was over 30 per cent.

booms and slumps less intense, and help to level out fluctuations in trade and employment.

Objections are sometimes brought against such a policy on grounds of finance. It is suggested that public works designed to absorb the unemployed are needlessly expensive, and that the State will have difficulty in raising the necessary capital, or alternatively will run too rapidly into debt. The first objection is based upon a misunderstanding. The public works programme which it is proposed to carry out in times of depression is not drawn up for the purpose of *creating* work; it is not a programme of *relief* works. It includes —or should include—only schemes which the government would wish to carry out anyhow,[1] and which create vacancies for men in their ordinary employment at the usual rates of wages.[2] So far from raising costs a policy of transference of public works from boom to slump will avoid the high costs of constructional work at the peak of the boom and effect economies at the lower prices ruling in the slump. Nor is the debt problem likely to be an important stumbling-block. The suggestion that savings will be inadequate in times of slump confuses cause and effect. Savings are low because income is low; and income is low because investment is low. If investment is increased, therefore, income will rise simultaneously, and the savings necessary to finance the new investment will be automatically forthcoming out of the rise in income. It is, of course, possible that the government's credit may be so much worse during a depression that it will require to offer higher rates of interest in order to procure loan capital. But this is most unlikely in the early stages of a depression when interest rates are falling; and will only be true in the later stages if the government has let the slump get completely out of hand by its failure to adopt just such expedients as a policy of public works. As for the suggestion that the government may run into debt too rapidly, this is clearly not an objection to *postponement* of borrowing so long as the total amount borrowed remains unchanged. And even if more is in fact borrowed in order to reduce unemployment, much of the additional debt burden would have to be incurred anyhow—failing a spontaneous recovery in private industry—in order to meet the cost of unemployment insurance benefits and unemployment allowances. It is surely better to run into debt in order to enrich the country with useful assets than to run into debt with nothing better to show than high unemployment percentages.

The real difficulties are not financial. They are, first of all, diagnostic. Can we rely upon governments to decide on the right moment to put

[1] In practice, of course, a government may be induced to expand its programme of construction beyond what it would have envisaged in the absence of unemployment: these *additional* schemes require to be justified on other grounds.

[2] The employment created will be in a wide variety of trades, since men will be set to work to make *materials* for the constructional work undertaken, and since there will be an increased demand for *consumption goods* out of the larger incomes of those who find employment.

into operation schemes of public works prepared beforehand? May they not be tempted to continue a programme of construction into the boom, through failure to read the signs of a slump to come? Or even to accelerate their programme when prosperity and budget surpluses spread a mood of optimism? May they not, by premature expansion of investment at the first signs of recession, simply restore and intensify boom conditions? Or, in a depression, may they not confuse with normal cyclical unemployment, unemployment which is due to structural changes in industry or to transcyclical shifts in the location of industry, and apply a remedy which, while appropriate to the one, is not necessarily appropriate to the other?

These difficulties are aggravated, at any rate in Great Britain, by political factors. It is much easier to bring pressure to bear in favour of an expansion than in favour of a contraction in public works expenditure; so that, in practice, there is generally a steady expansion, interrupted by panic curtailments such as took place in 1931, instead of a contra-cyclical oscillation between boom and slump. The fact, too, that British budgets are drawn up on a yearly basis, and make no attempt to distinguish carefully between income and capital items, puts special difficulties in the way of a consistent programme of public works stretching over the period of the trade cycle. A budget which provides for heavy borrowing in order to finance public works in a depression can be represented as "unbalanced" (a term of abuse with no clear meaning); and a programme covering a long period of years can be repudiated by a new government if it so chooses.[1]

A third difficulty is the "intractability" of public investment. Some kinds of construction simply cannot be postponed (for example, the building of warships or the replacement of condemned bridges); others can be postponed only at the risk of a public outcry (slum clearance, for example); others, again, must, once commenced, be carried through to completion, boom or no boom (major schemes of road improvement, bridge building, the construction of docks, harbours, electric power stations, and so on). Moreover, it is not always easy for a local authority or a central government to foresee its requirements over the next five or ten years, since new duties may be imposed upon it, new types of equipment may become available, and changes may take place in the size and kind of population for which it has to provide. The preparation of plans in advance, therefore presents great difficulties unless the government, local or central, holds itself free to make extensive departures from the original plans.

It is possible also that a policy of public works may fail in its object because its effects are discounted by private enterprise. If business men look on public works as a kind of pick-me-up, creating an "artificial" and temporary prosperity, they will be deterred by fear

[1] In Sweden, the practice of having an extraordinary budget, covering a period of five years, makes it much easier to increase loan-expenditure at one time and make repayments of capital at another.

of the ensuing "headache" from reacting with the same vigour that a so-called "genuine" and spontaneous revival would produce. They will be induced to withhold from the private investment which should replace public works, and which public works are intended to underpin, through fear of a fresh collapse of industrial activity immediately the props which support it are withdrawn. Such perversity, however, is to be expected only in special circumstances—when government policy is half-hearted or does not command the confidence of business men; and when the marginal productivity of capital is so low that the scope for profitable investment by private enterprise is seriously limited in relation to potential savings.

A policy of large-scale public works is of doubtful value when depression is centred mainly in a country's export trades. The employment provided must be as far as possible a substitute for the employment lost, and it is comparatively rare for a large proportion of the workers in the export trades to be readily adaptable to the requirements of domestic construction. They may, however, be comparatively well suited to supplying the domestic market with *consumption* goods, so that a Socialist government which did not seek to recover the full cost of such goods out of sales proceeds would not find itself in quite the same *impasse* as a Capitalist government. Generally speaking, however, whatever the form of government, any country which is highly dependent upon foreign trade will have more difficulty in avoiding unemployment than a country with a large domestic market.

The limits to a policy of public works are comparatively narrow. A capitalist government can rarely undertake the manufacture of consumption goods, since it would enter into competition with private enterprise; and the forms of construction for which it is responsible may be inadequate to maintain investment at its optimum level. On the one hand, therefore, a capitalist government is limited to the construction of capital assets which may be of little utility, and is prohibited from using the unemployed in order to raise the general standard of living by increasing the output of essential consumption goods; on the other hand, it may find itself powerless to cure a chronic and severe depression because of the limited scope of the constructional activity in which it can indulge. It is a strange paradox that the one kind of construction in which saturation is never reached, and in which the "legitimate sphere" of private enterprise is not infringed, is the manufacture of instruments of destruction. The one form of "public works" which cannot fail to cure unemployment is war and preparation for war.

There is little doubt that, when other ingredients making for recovery have been present, a policy of public works has been effective in speeding-up that recovery. There is no doubt, too, that when we have to deal only with cyclical unemployment, and with a regular alternation of mild booms and equally mild slumps, a policy of public works can make booms and slumps still milder. But whether such a policy is by itself an effective remedy for long-

term unemployment in declining industries, or for a chronic and prolonged depression over the whole field of industry, is open to question.

**(ii) Credit Policy and Control of Investment.**—The State may seek to influence the level of investment by more indirect expedients. It may, by a suitable monetary policy, try to bring down interest rates in the slump and raise them in the boom. In this way, industry will be offered an inducement to increase capital outlay at times when it is low and to restrict it at times when it is excessive. At the same time, the State may set up a National Investment Board to supervise new capital issues and prevent some of those misdirections of investment which disfigure the boom and, by destroying confidence, aggravate the subsequent slump.

The State has comparatively little *direct* control over long-term rates of interest. It can bring pressure upon the banking system to buy or sell bonds; and it can, by regulating the conditions under which new issues are made on the Stock Exchange (for example, by excluding foreign issues) make borrowing easier or more difficult. It has also control over various extra-budgetary funds which are administered through the National Debt Commissioners, the Unemployment Insurance Committee, and similar public bodies; and by varying the securities in which these funds are invested (for example, by switching from short-dated to long-dated securities), it can, within narrow limits, change the structures of interest rates. But in the main, the State is forced to influence long-term rates *indirectly* by Bank Rate policy. Now Bank Rate policy, as we have seen, generally affects only short-term rates of interest, and even a very large change in short-term rates will be ineffective in producing a comparable change in long-term rates, particularly if there is a general belief that the change in short-term rates is temporary and that it will be reversed within a few years or even a few months.[1] In the boom, a slight rise in short-term rates may take effect because business men interpret it as a signal of the central bank's intentions and know that it can be followed up, if necessary, by a much larger rise. But in the slump, there is a clear limit to the reduction in short-term rates which the central bank can bring about, and an equally marked limit, therefore, to the effectiveness of monetary policy. The water of cheap money may be spread as abundantly as we choose but industry may disdain it. Indeed, the water may be so shallow that industry would have the greatest difficulty in drinking. Low short-term rates are not likely by themselves to encourage additional borrowing since the demand for short-term capital does not greatly depend upon its price.[2]

[1] *cf.* above, page 259.

[2] A few economists, notably Mr. Hawtrey, hold a contrary view, and believe that the demand for bank advances (with a view to financing the holding of stocks) is sensitive to changes in the terms of credit. Mr. Hawtrey, therefore, has no reason to be deterred from urging monetary policy as a cure for unemployment by his conviction that short and long rates of interest have little influence on one another.

Cheap money can encourage industrial borrowers only through its influence on the price or marketability of industrial debentures and equity stocks. But these are even less sensitive to changes in short-term rates than are government bonds. Credit policy in a slump, therefore, is not a very effective instrument in promoting recovery, except over a period of years or in conjunction with a policy of public works. Similarly a National Investment Board would be of much more use under boom conditions in influencing the *quality* of investment than in a slump when the quantity of investment would be of prime importance. If the National Investment Board was charged with the supervision of a public works policy, however, its functions would clearly be less negative and its power to influence the volume of investment enormously greater.

**(iii) A Reduction in Working Hours.**—One way of dealing with unemployment is to spread work—to turn out the same output by employing a given number of workers in rotation or by taking on more men and working a shorter day or a shorter week. Neither of these devices, however, is altogether satisfactory. The first spreads unemployment just as fast as it spreads work—some men lose a day's work in order to provide work for others. Not only is such a scheme difficult to operate; it is apt to prolong the depression which it aims at overcoming since by offering *all* workers a chance of employment in contracting industries it deters *some* of them from moving to expanding industries under pressure of unemployment. Even in a cyclical depression when there are no expanding industries to move to, the spreading of unemployment is not likely to be of great value. Unemployment will be concentrated in particular industries and areas, so that the burden will continue to be borne very unequally by different groups of workers in spite of spreading, while the constant change in the working personnel will go far to ruin efficient production. A reduction in working hours raises fewer administrative difficulties, and has the advantage of converting unemployment into genuine leisure, not periodic and enforced idleness. It is probably a more suitable method, therefore, of dealing with chronic unemployment in industries in which labour is particularly immobile (e.g., the coal industry). But it is not a very satisfactory method of countering a normal cyclical depression unless in some mysterious way it restores the incentive to invest.

The first effect of a reduction in working hours is generally to raise wage costs. Hourly wage rates necessarily rise if, as is usual, the reduction in hours is not accompanied by a corresponding reduction in weekly wage rates. There may be a small increase in the worker's efficiency, or a slight improvement in methods of operation,[1] but these are unlikely to offset the rise in money wage rates.

[1] By stopping work on Saturday morning and working a five-day week, for example, a firm may cut out the most costly part of its output: so that even a slight reduction in weekly wages might compensate for the loss of product.

The second effect of the reduction in hours is to raise prices and the cost of living. Employers will seek to recoup their higher costs from higher prices, and will be successful in raising prices more or less in proportion to the rise in marginal prime cost, unless consumers show a disposition to curtail purchases and accept a lower standard of living.[1] It is likely that some classes in the community (those, for example, with fixed money incomes) will be unable, at the higher level of prices, to maintain their purchases. Demand, therefore, will fall off; and with demand, output and prices. The fall in output, thirdly, will reduce the volume of work to be done. But this may nevertheless be quite compatible with a decline in unemployment. For if hours of work have fallen by, say, 20 per cent., and the volume of output contracts by, say, 12 per cent., there will be an increase in employment (initially at any rate), by 10 per cent. The more tenacious people are of their old standard of living, the greater are the chances of an expansion in employment. This is equivalent to saying that unemployment will disappear the bigger the change in thriftiness, for at the higher level of costs it is only at the expense of thrift that the old standard of living can be maintained.

Reducing working hours, therefore, is a suitable expedient for curing general unemployment only when a fundamental change in habits of thrift is called for, and when there is no more effective way of bringing about such a change. It is a risky expedient, since it may increase, not diminish, unemployment—either by failing to reduce thrift sufficiently; or by raising costs in industries which supply foreign markets or are unsheltered from foreign competition; or by producing an abrupt rise in prices which destroys the confidence of investors and deters business men from contemplated loan expenditure. It forces workers to choose leisure when they might prefer fuller employment and a higher pay with working hours unchanged. And it does not even, as is popularly supposed, redistribute income to any great extent in favour of wage-earners. For, although money wages may not be reduced, the subsequent rise in the cost of living eats into real wages. *Hourly* real wages are higher than before, but not *weekly* real wages.

**(iv) Wage-cuts.**—The traditional remedy for unemployment is to reduce costs and the easiest way to reduce costs is to cut wages. It seems obvious that a reduction in wages must bring about an expansion in employment, for lower wages mean cheaper labour, and it is usual for demand to increase as things become cheaper. What is true of a single industry, however, may not be true of all taken together, for the costs of one industry are the incomes which go to buy the products of the others. If the wages of coal-miners fall,

[1] If the old ratio between price and marginal cost is restored, producers will have an incentive to market the same output as before; and if consumers, at the new level of prices, are willing to buy the same quantity of goods as before, the goods produced will find a market without change of price.

employment in the mining industry will increase; but since coal-miners will have less purchasing power, employment in other industries will decrease. We cannot say, *a priori*, whether the wage cut will increase or diminish employment.

The same conclusion may be reached by a different route. If money wage rates are reduced in a given proportion, there are good reasons for expecting competition to force down prices in much the same proportion. But if prices and wages fall equally, *real* wages will be unchanged and there will be no incentive to employers to expand output. It may be extremely difficult to reduce real wages by all-round reductions in money wages; and changes in money wages, therefore, are likely to be comparatively ineffective in influencing the volume of employment.[1]

Although *all-round* reductions may do little good, reductions in selected industries where the demand for labour is particularly elastic may accomplish a substantial net increase in employment. It may happen, for example, that prices and wages in one of the constructional trades are widely regarded as abnormally high, and that such high prices are a deterrent to investment. A wage cut, by bringing prices into line with accepted norms, may produce a wave of buying orders which make a real contribution towards recovery. Wage cuts in some of the export industries, supplying extensive foreign markets, may also be effective in increasing employment.

In general, however, wage cuts are a competitive device to allow one business or industry to steal a march on others, not an adequate remedy for slumps. Of course, if it is felt that no recovery can begin without wage cuts, then no recovery *will* begin without wage cuts; for the general opinion, by suspending investment, will intensify the depression and so provide its own corroboration; just as, once wage cuts *have* taken place, and *have* encouraged new investment, the subsequent recovery will be taken as proof that the general opinion was right. Whatever myths receive general acceptance amongst investors will be transmuted by the mere fact of their acceptance into sober truth; and whatever sacrifices are required on the altar of investors' beliefs will appear just and necessary—not the fruits of delusion.

(v) **Other Expedients.**—Of the other expedients which have been suggested as remedies for unemployment, without involving a fundamental change in the economic system, few are worth discussing in detail. One—protection—has already been examined above.[2] An alternative to protection is the use of subsidies. By granting subsidies in times of depression and removing them in times of boom, the State might alternately increase and decrease the profit margin

[1] As was argued above (page 250), it is chiefly through changes in the rates of interest and profit that real wages are affected, not through changes in money wages.

[2] See page 289.

according as it showed signs of becoming insufficient or excessive for the purpose of maintaining full employment. Subsidies, too, could be devised which would give the maximum of encouragement to investment and employment. A subsidy on house building, for example, would exert the maximum pressure on construction if a rigid time limit were set, since builders would speed up work in hand so as to qualify for the subsidy. Or the subsidy might be made to vary between different parts of the country so as to offer the maximum incentive to construction in districts where the demand for houses was elastic.

There are clear limits to the uses to which subsidies can be put. First of all, since a subsidy is paid out of taxation, it is necessary to raise additional taxes which may depress industry almost as much as the subsidy stimulates it. Secondly, just as protection, once introduced, tends to stay through boom and slump, so subsidies, once granted, tend to continue indefinitely. Thirdly, subsidies are a half-way stage to industrial bankruptcy with the State as receiver. If private enterprise cannot pay its way, there is no reason why the State should maintain it on the taxpayers' charity, and the natural sequel to a State subsidy is State control. This sequel may be long deferred, however, especially if industry shows strong powers of recuperation, or if there is a deep-seated prejudice against nationalisation in any form.

The State may also adopt a policy of deliberately encouraging monopoly and restriction. It may hold out to entrepreneurs the prospect of higher prices and profits through a concerted restriction of output. In industry, firms will be freed from anti-trust legislation and invited to co-operate with one another in taking joint action against the depression (in other words, in raising prices); in agriculture, farmers will be paid to restrict the area sown, or to destroy growing crops, or to kill off livestock. All these expedients were used in the initial stages of President Roosevelt's New Deal policy. But as remedies for trade depression they are manifestly absurd.[1] For if everyone is busy *restricting* output in order to raise prices, there can be no *expansion* in aggregate output and employment. It is true that prices generally rise as recovery proceeds, but this is a symptom of *increasing demand*. To aim at a rise in prices for its own sake, even at the cost of *decreasing supply*, is to cut off an ugly nose to improve one's looks. We do not increase output by seeking to make things scarcer, nor get rid of unemployment by causing more.

Many other remedies for unemployment—raising the school-leaving age, pensions at 60, and so on—might be mentioned, but they raise few novel points of principle, and their merits can probably be judged on the basis of the above analysis. We can pass on, therefore, to a third, and final, deficiency of laissez-faire.

[1] There was something to be said for a policy of restriction where a particular product had been overproduced even in the boom of 1929. But the actual policy adopted went far beyond this.

**(c) Laissez-faire does not solve the Problem of Poverty and Inequality.**—If every man is free to earn what he can, in a world which offers large profits to the enterprising and the fortunate, it is to be expected that some people will earn much more than others. These inequalities, especially when they are clearly due to luck or inheritance, and when the many still go in fear of unemployment and poverty, arouse envy and resentment; in a democratic community, where everyone is equal in voting power, it is felt to be inconsistent to permit wide differences in purchasing power.

But how are these wide differences to be prevented? One way would be to adopt a communist system, in which everyone would be employed by, and paid by, the State, and in which all (or nearly all) earnings would come from labour since all (or nearly all) property would be publicly owned. As the sole employer of labour, the State would be able to keep inequality of earnings within prescribed limits. This is a solution more drastic than this country has so far been prepared to face. It may be doubted whether, if all that was in question was inequality of income, the expedient of communism would be much more successful than the measures adopted in Great Britain; wage rates in Russia in 1941 varied almost as enormously as did British incomes (after payment of tax) in the same year. There is, moreover, a deep-seated distrust in Great Britain of anything that concentrates power too exclusively in the hands of the central government, however exalted the motives for such concentration. A universal paymaster may be a universal despot.

**The Liberal Compromise.**

There is a simpler way in which the State can make its influence felt to check inequality. It can play Robin Hood by taking from the rich and giving to the poor. By heavy death duties the State can prevent the passing of large estates intact to private persons by the accident of birth or favour; by steeply progressive rates of taxation[1] it can keep earnings, after payment of tax, within relatively narrow limits; by creating a system of social services, it can guarantee a minimum standard of living. All this within a social framework in which the flexibility of private enterprise can still, in great measure, be preserved. It is just such a compromise that was in process of being worked out in Great Britain before the war; and if the compromise was not by any means to everybody's liking, it was attacked on grounds of inequality of employment rather than of income; or because, in spite of agreement on the principle, the scale of transfer from rich to poor was thought far too modest.

**(a) Progressive Taxation.**

The "Liberal Compromise" (as we may term it) involved a revolution in the theory of taxation. In 1900 it was almost economic heresy to suggest the deliberate use of a progressive income tax to procure greater equality of income. Economists defended progressive

[1] That is, rates which are higher, the richer the taxpayer.

taxation on other grounds.[1] They cited the rather oracular pronouncement of Adam Smith that people "ought to contribute towards the support of the government, as nearly as possible in proportion to their respective abilities," and undertook to prove that ability to pay increased more than in proportion to income. Or they put forward an "equality of sacrifice" theory and engaged in philosophic discussions on the rate at which the income tax would require to increase as income increased in order to involve rich and poor in an equal measure of sacrifice.

Whether their slogan was "ability to pay" or "equality of sacrifice," economists found arguments to justify progressive taxation. But they did not regard it as part of a comprehensive policy for checking the growth of large fortunes and financing a system of social services. The theory of progressive taxation was rather like that of discriminating monopoly—mathematical rather than ethical. The question posed was: Given the revenue to be raised and the number and relative incomes of the taxpayers, how much should each pay in tax? But there is, of course, a more fundamental question to which this leads: What incomes should people be left with after the State has raised and spent its revenue?

Our answer to this question depends upon the importance which we attach to ability and need as alternative principles of distribution. Men are both producers and consumers; and the distribution of income which spurs them to the greatest effort as producers may be grossly unfair when we look at the needs which have to be met out of the unequal earnings. But whether rewards should be related to ability and effort or to human needs, neither principle can justify the prevailing extremes of poverty and wealth. Rewards are obviously not in accord with needs. Neither are they proportioned uniformly to ability and effort. They depend also upon the wealth or opportunities bequeathed by one's parents; the efforts which are rewarded do not always confer a benefit on society, but may be directed to anti-social ends; and it is highly unlikely that the highest incomes are a necessary spur to the efforts and sacrifices for which these incomes are paid. It is possible, therefore, for a progressive income tax simultaneously to bring income and need, and income and effort, closer together. Moreover, the incentive that controls effort is frequently gross, not net, income; we are often content with the shadow of purchasing power while the government makes off with the substance. Hence taxation may go very far towards lining up income and needs without doing damage to the relationship between income and effort.

[1] Mention may be made in passing of a once popular theory of taxation which was used in attacks on progressive taxation: the so-called "benefit" theory, which suggested that taxes are a payment for services rendered by the State. This theory naturally recommended itself to those who disliked taxes on principle and looked on the State as a kind of grocer's shop selling law and order to those who stood to gain from the purchase.

But not indefinitely far. There are limits to the usefulness of taxation as a means of redistributing income. In the first place, a high rate of tax affects people's inclination to work, and hence the supply of labour. Since the highest rate of tax is paid on marginal earnings, there is a natural temptation to work less hard, do less overtime, retire earlier from business. A skilled workman or professional man who is paying income tax at 10s. in the £1 on part of his earnings will naturally reflect that every £1 that he is offered for overtime work is worth only 10s. net of tax. If there is a fixed net income which he is determined to earn, he will have to work all the harder because of the tax which he has to pay. But he will generally be inclined to react the other way and sacrifice the 10s. for the sake of some leisure. After all, it is because he is fairly well off that he is liable for income tax at so high a rate; and since he is well off he is likely to be less rigid in his idea of the income that he wants to earn.

High taxation reacts on savings in a similar way. It reduces the surplus out of which people save; and it penalises saving more than expenditure, since the income set aside to be saved is taxed at the time, and when dividends are paid on the capital sum they, too, are taxed. On the other hand, when people are determined to make provision for a given *net* income from investments at some future time, the higher the rate of tax the more they will be forced to set aside. It is not a *necessary* consequence of high taxation, therefore, that it should reduce private savings; but we are pretty safe in concluding that in practice it does, on balance, have this effect.

This is of special importance when we are thinking of business profits. The traditional method of financing extensions to a successful small business is out of profit; but if a large part of the profit is drained away in tax, growth is retarded and the attractiveness of founding a small business is greatly reduced. One effect of a high income tax, therefore, is to fortify established businesses against competition from young and expanding rivals, and to alleviate one of the evils of private enterprise (poverty) by aggravating another (monopoly).

An additional reason why this tends to come about is that high taxation penalises risk-bearing. What was previously an even chance is now more likely to be a loss since tax is deducted from any profit earned, but no compensating refund is made for a loss. If, for example, there was formerly an equal chance of a profit of 10 per cent. or a loss of 2 per cent., the effect of an income tax of 10s. in the £1 is to turn this into an equal chance of a net profit of 5 per cent. or a net loss of 2 per cent.—obviously a much less attractive proposition. The established concern whose risks are limited gains once again by comparison with the young and struggling firm pioneering products or methods which have yet to be tested. The check on enterprise has also its repercussions on employment and, by reducing the resiliency of the economic system, tends to prolong periods of slump.

These limitations to taxation are obviously elastic and can be largely offset by other measures. If private savings diminish, the State can use public funds for capital construction. If the enterprise of young firms is checked, the State can foster new methods of industrial finance designed to put capital at the disposal of new and expanding businesses. If there is chronic unemployment, there are better ways of dealing with it than by revising rates of income tax. And if high rates of tax cause earlier retirals from business and less overtime, who is to say that the community (except in war-time) loses by the change ?[1]

Whatever the theory of the matter, we have moved, within 50 years, from a state of things in which the largest incomes did not pay more than about 5 per cent. in tax and the lowest incomes paid at least as much, to one in which the largest incomes pay over 90 per cent. in tax and the lowest incomes receive much more from State subsidies and services than is taken in tax. Fifty years ago there was no limit to the accumulation of private fortunes ; now it is practically impossible to maintain intact and pass to one's heirs more than about £100,000, and practically impossible to earn an income, net of tax, of more than £5,000. These are upper limits ; and, if current trends continue, and we ignore a few scores of taxpayers, the limit to net incomes will soon be nearer £2,000 to £3,000.

This change has been associated with the change in the electorate from the middle-class monopoly of the nineteenth century to the universal suffrage of the twentieth ; and it has been greatly accelerated by the three costly wars which we have waged since the century began. It has been brought about, not by Socialist governments pledged to " soak the rich," but mainly by Conservative and Coalition governments in need of enormous revenues and anxious not to antagonise the wage-earner.

**(b) The Minimum Standard.**

The State is concerned not only to prevent the accumulation of vast economic power in private hands, but also to get rid of poverty by establishing a minimum standard of living for all citizens. This it tries to accomplish by means of what have come to be called " the social services " ; by schemes of insurance against ill-health and unemployment ; by pension schemes for widows and the aged ; by housing subsidies to enable additional houses to be provided at reasonable rents ; by food subsidies to keep down the cost of important foodstuffs such as milk, and to make it easier for school-children, nursing and expectant mothers, and other " priority " consumers to be properly fed ; and by a host of other services, the aim of which is by no means simply to relieve poverty, but is rather

[1] There will, however, always be some ambitious men in business who are held back from exploiting new inventions by the absence of a profit inducement. A man who can easily earn £1,000 a year will not submit himself to a tremendous strain for a year or two merely for the sake of a few thousand pounds more.

to extend to everyone on a communal basis some of the security and amenities that were once the privilege of the rich.

Expenditure on the social services has increased rapidly since the beginning of the century. In 1910 the total was £63 millions, and in 1936, excluding war pensions, £455 millions. How these totals were distributed between the different services is shown in the following table:—

EXPENDITURE ON THE PUBLIC SOCIAL SERVICES IN 1936.

| | £m. |
|---|---|
| Unemployment Benefit and Allowances | 86·3 |
| Poor Relief | 51·5 |
| Widows', Orphans' and Old Age Pensions | 90·1 |
| Education | 116·0 |
| Housing | 43·8 |
| Health | 67·0 |
| | 454·7 |

It must not be imagined that the entire cost of these services is borne by the general taxpayer. Much of the cost is met from contributions by the workers who benefit from the expenditure. In 1936, for example, £150 millions came from contributions, fees, rents, etc., and the balance of £315 millions from the Exchequer and from local rates.[1] The total of £465 millions included £80 millions paid by employers and employed under the Health and Unemployment Insurance Acts, £23 millions paid in rent by the occupants of houses in public ownership, and £31 millions paid in contributions towards State contributory pensions schemes. It is obvious, therefore, that the net transfer of income to wage-earners through State expenditure on the social services falls substantially short of the gross outlay of £455 millions and is probably about half that figure.

One great merit of the social services is that they confer benefit without the stigma and humiliation of charity. Benefit is drawn as of right. The State lays down the terms upon which the service is made available and, so long as those terms are complied with, anyone may apply. In this there is an obvious danger of demoralisation. If people acquire the habit of applying to the State as fairy godmother, they may spare themselves the trouble of fending for themselves and cast themselves without compunction on the indefinite charity of their neighbours; they may even conclude that if so much can be given so freely, more might be given in response to pressure. This danger, however, is more obvious than real. No one nowadays would think of urging it against free education, or health insurance, or even the milk-in-schools scheme. It is commonly urged

[1] Total receipts from all sources in 1936 were about £10 millions in excess of expenditure.

(on rather slender evidence) against high rates of unemployment benefit. As a rule, however, it is put forward in quite general terms. And on general grounds we should probably get nowhere if the only assistance that we were prepared to offer the poor was assistance that could be guaranteed not to demoralise, and that could not honourably be accepted.

A second merit of the social services is that they are on such a large scale that they are exceptionally cheap to administer. The economy of State insurance and pension schemes is particularly striking. It is not uncommon for even a fairly large insurance company to find that nearly half of what it receives in premiums is swallowed up in the cost of administration. But the cost of administering the Unemployment Insurance scheme is well below 10 per cent. of the insurance contributions. There are good reasons for expecting an extension in State insurance (e.g., to burial insurance) to yield economies at least as striking. It is also to be expected that in larger countries, such as Russia and the U.S.A., economies in administration will be even greater than in Great Britain.

Thirdly, the kind of service which the State wishes its citizens to have may involve such a close degree of State supervision that no private body could be relied upon to provide it.

**Summary: Planning and Power.**

In the economic order as we know it, the planning undertaken by consumers in deciding upon their budget, and by producers in deciding upon their production programme, is supplemented by planning by government officials. This planning is necessary because the plans of individuals are not sufficiently comprehensive and may result in monopoly, unemployment, poverty and other evils. But it differs from the planning of individuals in that it rests, not upon incentives as expressed in prices and profits, but upon compulsion backed by the force of law. In practice, the contrast is less sharp because plans do not pass into law and are not framed by government officials under a democracy unless they command a wide measure of approval. But in principle the distinction between State planning and what has been referred to in this chapter as "laissez-faire," is clear; State planning leaves the last word to a government official, while under "laissez-faire" decisions are taken without such intervention.

Reliance upon price incentives—whether profit-motive or wage-motive—is often disparaged unthinkingly by those who have no grasp of the function of those incentives, and suffer from a cloudy aspiration after ideals of service and the direction of economic planning by some benevolent master-mind at the centre. The plain fact is that master-minds capable of such direction do not exist; or, if they do, they are not at the centre; or, if they are, they are not benevolent, or not benevolent for long. Planning from the centre means creating a most complicated piece of machinery which can never be properly oiled; it can never, viewed as a piece of machinery, be as efficient

as the planning that goes on within a small firm. Planning from the centre means, secondly, the concentration of power to take decisions at the centre; and this carries with it a threat to democratic rights. What is initially the subordination of individual preferences to an order and organisation conceived in the interests of the community, may develop into a denial of political rights; those who enjoy the power of decision over economic organisation may not relish the curbs and checks of popular government, and may resort increasingly to force to perpetuate their power or to implement their planning. Private enterprise, whatever its failings, at least disperses power and initiative.

In the nineteenth century, when liberty was the loudest cry, this line of thought carried the day. But in the twentieth century, when people call most for security, order, equality and so on, some compromise will obviously be necessary. For security, order and equality, unlike liberty, are not things that just happen because the State stands aside; they have to be contrived, and it is in the contriving that the danger lies. Contriving means creating a bureaucracy; and a bureaucracy may become a caste, and take to itself a dictatorship.

Neither State planning nor private enterprise is without blemish. On the one hand lie:—

> "The law's delays,
> The insolence of office and the spurns
> That patient merit of the unworthy takes."

On the other, is the degradation of unemployment and poverty, and the frustration of initiative through monopoly. It will be a major task of statesmanship for the rest of this century to evolve an economic order which eliminates most of the wastes of private enterprise without putting on the State a burden greater than it can safely carry.

There is no blue-print for such an order; it should grow out of the experience which we accumulate in war and in peace. Economists may sift that experience, sort out the problems still to be solved, and demonstrate how much those problems have in common with one another. But they are unlikely to agree on some solvent by which all of them can be made to vanish. All that they can do is to inculcate the right temper and technique of thought with which to approach the problems, and by skilful diagnosis make hasty and inappropriate remedies give place to the wisdom of mature understanding.

Acceptances, 308-9, 312
Acceptance Houses, 358
Accountancy, 68, 118*n*, 174*n*, 204 ; and economics, 8 ; and scale of production, 81
Advantages : natural and comparative, chap. 5 *passim*, 7, 44, 281-2 ; net, 237
Advertising, 74, 77, 109, 118*n*, 403 ; and choice, 5 ; and scale of production, 85 *et seq.* ; by industries, 178
Agriculture, 27, 30-2, 54, 64, 66, 76, 79, 133, 135, 189, 195, 203-4, 214, 220-1, 249, 284, 313 ; and margin of transference, 185-6 ; and rising costs, 191 ; and trade cycle, 382 ; wages in, 225, 240-1
Allocation : of resources, 4 ; *See also* Rationing
Allowances : family, 376
Aluminium, 132, 187
Amenities : and wages, 238
Artificial Silk, 23
Astrology, 12

Baking, 84
Balance : of payments, 331, 363 ; of trade, 37, 280, 285, 398-9
Bank : advances, 74-5, 259, 313-5 ; deposits, 309, 346-7 ; notes, 319 *et seq.* ; rate, 313-4, 324-5, 382, 410 ; reserves, 310
Bank Charter Act, 319-21, 365*n*
Banking : branch, 317 ; central, 304, 319, 360 *et seq.* ; school, 320
Bank of England, 186, 294, 298 *et seq.*, 303-4, 319 *et seq.*, 329 *e seq.*, 362 *et seq.*
Banks, 42, 43, 68, 96, 125, 235, chap. 24 *passim* ; defined, 302 ; country, 304 ; City of Glasgow, 90*n*, 307 ; French, 317-9 ; German, 317-8 ; joint-stock, 306-7 ; London, 303 ; private, 306 ; Scottish, 307 ; U.S., 179, 317-9, 330*n* ; U.S. amalgamation of, 306-7
Barter, 1, 292
Berle, A. A., 97*n*, 102, 103
Bills of Exchange, 311-2 ; Inland, 304-5, 311, 362*n* ; Treasury, 312
Bill-brokers, 305, 318-9, 363, 366
Bishirgian Pepper Corner, 100*n*
Blacklegs, 249
Blum Experiment, 41, 250
Booms and Slumps, chap. 28 *passim*, 76, 101, 129, 316, 344, 406
Bootmaking, 58
Bottlenecks, 68, 84, 379
Boulder Dam, 47
Bowley, Professor A. L., 20
Branch Factories, 86, 107, 127, 134 ; banking, 307
Bresciani-Turroni, C., 351*n*
British Broadcasting Corporation, 143
Budget, 20, 100*n*, 410
Building, 34, 57, 190, 192
Building Societies, 96, 260
Bullion, 360
Bullion Report of 1810, 370*n*
Bulls and Bears, 99
Bureaucracy, 4, 421
Burns, A. R., 104*n*, 132*n*

Cambridge Formula, 2
Campbell, C. D., 100*n*
Campion, H., 17, 140*n*
Canals, 89
Capital, chaps. 8 and 21 *passim* ; accumulation, 43 ; as a factor of production, 22 ; concrete capital, 22, 35, 93, 383 ; debt capital, 39-40 ; definition of, 35 ; function of, 262-5 ; trade, 35 ; loan or money capital, 22, 36, 325-6, 356 ; longevity of, 93 ; marginal productivity of, 265-7 ; mobility of, 261 *et seq.* ; market, 95-6, 260, 280, 295 ; outlay, 19, 354-5, 382-5 ; social capital, 35 ; *See also* Interest, Investment, Profit, Savings
Capitalism, chaps. 10 and 30 *passim*, 210-1, 257
Cartels, 121, 183
Cash Credit Bond, 308
Cassel, Professor Gustav, 370*n*
Chamberlin, E. H., 113*n*
Cheques, introduction of, 298*n*
Choice : and planning, 138 ; and rationality, 5 ; freedom of, 3, 6, 212 ; *See also* Consumers' Preferences
Christmas Presents, 8

Clark, Colin, 20, 97*n*
Clark, J. M., 66*n*
Clarke, Tom, 83
Class Divisions, 15, 140-2, 243-4, 254
Clothing Trades, 59, 111
Coal-mining, 11, 16, 23, 54-5, 67, 121, 124*n*, 151, 184, 189, 210-1, 214, 240-1, 377-8
Coefficients, technical, 228
Coins, 296-7
Cole, G. D. H., 381*n*
Collateral, 308
Collective Bargaining, 232-3
Colonisation, 43, 267, 399, 401
Colour-printing, 54
Combination, industrial, chap. 9 *passim*, 64
Commodity: difficulty of defining 154, 177
Communism, 415
Competition, Part 3, *passim*; defined, 177; in services, 181; perfect and imperfect, 23, 146-7, 177, 179-80, 188, 233, 266; *See also* Monopoly
Consols, 315
Consumers' Preferences, 4, 6, 167 *et seq.*, 180-1, 212-3
Consumers' Surplus, 163
Consumption: Definition, 22: *See* Demand
Co-operative Movement, 142
Concentration: of economic power, 101-3, 118 *et seq.*, 140-1, 420-1
Corners, 100, 140
Cost Curves, 155*n*
Cost of Living, 20, 37, 42. *See also* Index Numbers
Cost of Production: and alternatives, 29, 189, 207, 219; and rent, 29, 190; and value, 3, 15, Part 3 *passim*; in relation to output, 34, 88, 106-9, 175; in relation to time, chap. 16 *passim*; average and marginal, 172 *et seq.*, 204; prime, 199; shutdown, 201; transfer, 220
Costs: of entry, 59, 112, 172, 186; of transport, 56; joint, 150; overhead, 30, 198-200; real, 218; selling, 72, 172, 182, 206
Cotton Industry, 1, 23, 45, 49, 54, 57, 81, 120, 128*n*, 131
Credit: creation of, 42, 314-5, 324 *et seq.*, 363-5: *See also* Bank Advances
Currency, 37, 296-302
Currency and Bank Notes Act, 322
Currency School, 319-20
Current Account, 300-1
Custom: influence of, 5, 180, 234, 240, 255, 260, 296, 314, 354; and monetary policy, 330, 332

Daniels, Professor G. W., 140*n*
Death Duties, 415
Debentures, 93, 411
Debts: and money, 294-5; *See also* National Debt
Demand: chap. 13 *passim*; conditions, 151: curves, 152; deficiency of, 41-2; derived, 384; joint and composite, 202-3
Deposits: *See* Bank
Depreciation of currency, 30, 38, 198, 256, 261, 266, 294, 296-7, 374
Depressed Areas, 58
Depression, industrial, 380-3
Devaluation, 367-8, 374
Dewing, A. S., 74*n*, 85*n*, 117*n*, 124*n*, 125*n*
Disraeli, 51*n*, 100*n*
Douglas, Major, 381
Dictatorship, 4, 98; achievements of, 405; danger of, 421
Directors: of Joint-Stock Coys., 69-81-3, 90, 105-6, 140-2, 256, 277
Discount, 305*n*, 311-2, 324-5
Discrimination: price discrimination, 207-8, 211-2: in international trade, 285
Disintegration: vertical and lateral 57
Distribution of Income: Part 4, *passim*, 20, 27, 41, 146, 250; and fluctuations in prices, 355-6; and foreign trade, 290-1; and justice, 416 *et seq.*; and monopoly, 111-2; and unemployment, 382
Distribution of Property, 97, 140
Distributive Costs, *see* Advertising and Costs, Selling
Diversification of Output etc., 60, 76-8
Dividends, 91; *See also* Profits
Division of Labour, *See* Specialisation

Drugs, 14, 215

Earnings, *See* Wages, Income, etc.
Economic Man, 5
Economic Problem, 3
Economics : definition, 1, 7 ; and other subjects, chap. 1 *passim.*
Economies of Scale, chap. 15 *passim* ; external, 57, 61-4 ; internal, 61, 64 *et seq.*, 129
Edgeworth, F. Y., 166
Einstein, 243
Elasticity : of demand, 155, 233-4 ; of supply, 159 ; and substitution, 157, 226, 248; and time, chap. 16 *passim*
Electricity, *See* Power, 203
Electric Lamp Mfg., 45
Electric Washing Machines, 66
Emigration, 29
Employers' Associations, 245
Ends and means, 7
Engineering, 65, 87-8, 111
Entrepreneur, 405, 414 : *See also* Directors
Enterprise : as a factor of production, 22, 26 ; and location of industry, 58-9
Enterprise, Private, *See* Capitalism
Equilibrium, 152, 197 ; in Balance of Payments, 364
Establishment : definition, 61
Ethics, 1, 2, 4, 14-5
Exchange : and self-interest, 2 ; and specialisation, 45 ; implies interdependence, 2 ; rates, 360, 369-70, 373-4 ; triangular, 1 ; *See also* Barter
Exchange Equalisation Account, 322-3, 360, 368-9, 373
Exchanges : commodity, 24
Exploitation : of backward countries, 235 ; of labour, 262
Exports : chap. 22 *passim* ; after war of 1914-8, 377 ; invisible, 358

Factors of Production : chap. 3 *passim* ; and international trade, 286 *et seq.* ; earnings of, Part 4 *passim*
Feavearyear, A. E., 298*n*, 304*n*
Federal Reserve Board, 179, 319
Fens, draining of, 89
Fiduciary : note issue, 298, 320-3
Film production, 131
Finance : and size of firm, 74-5, 88-9, 95-6 ; absentee finance, 98 *et seq.* ; and industrial combination, 116 *et seq.* ; specialisation and finance, 43
Firm : definition, 61 ; as unit of control, 22 ; as independent unit, 118 *et seq.* ; co-ordination within, 50 ; growth biological, 127 ; and nation, 40-1 ; marginal, firm, 185, 201 ; optimum firm, 11, 68, 79, 80, 85, 106, 113, 123, 135 *et seq.*, 188 ; specialist firm, 63 ; cost curve of firm, 155*n* ; price policy of firm, 173 *et seq.*
Fisheries, 27, 263
Fisher, Professor Irving, 342
Fitzgerald, L., 123*n*
Florence, P. S., 82*n*, 87*n*, 129*n*
Fluctuations : in prices, 159, 169, 296 ; in trade and profit, 277 ; *See also* Booms and Slumps, Trade Cycle, Unemployment
Ford, Henry, 59*n*, 64, 69, 73, 82
Forward Exchange, 369
Furniture Trades, 84
Futures, 369

Gaitskell, Hugh, 351*n*
Gary Dinners, 119
Gilsonite, 60
Gladstone, 106
Glover, J., 54*n*
Gold-mining, 64, 179
Gold Standard : and money-supply, 37, 319-23, 326, 331, 360 *et seq.*, 382
Goldsmiths, 298, 303
Goods : capital, 38, 389 ; consumable, 38, 384-5, 389, 409 ; free, 3, 13, 18, 27 ; traded and domestic, 371
Goodwill, 15, 118*n*, 180-1
Graft, 100, 104
Grain milling, 59, 84
Grand Coulee Dam, 47
Grant, A. T. K., 96*n*, 97*n*
Gresham's Law, 303*n*
Growth : and integration, 127 *et seq.* ; biological, 127, 136 ; motives

to, 106 *et seq.* ; obstacles to, chap. 7 *passim*
Guns and butter, 10

Hammond, J. L. and B., 50*n*
Hardy, C. O., 75*n*
Hawtrey, R. G., 382, 404*n*, 410*n*
Hedging, 369
Hicks, Professor J. R., 20*n*
Hoarding, 346
Hobbies, 8
Hobson, J. A., 382
Holding Companies, 103 ; and liquidity, 316
Hot Money, 368
Hours of Labour, 230, 411-2
House-property, 17, 24, 36
Hume, David, 36, 285*n*, 402
Huxley, Aldous : on hair-dye, 213*n*

Ice-manufacturing, 23
Immigration, 227, 280, 288, 392
Imperial Chemical Industries, Ltd., 23, 71, 97, 111, 120, 122, 138
Imports, chap. 22 *passim* ; invisible, 350
Impulse, 7, 165
Incandescent Mantles, 23, 120
Incentives, 13 ; and size of firms, 82-3, 115 *et seq.*
Income : definition, 18 ; *See also* National Income
Income Tax, 415-7
Indebtedness : international, 359
Indeterminacy : in wage-rates, 245, 249
Index numbers, 333 *et seq.* ; Board of Trade, 339-40 ; Ministry of Labour, 386
Indifference curves, 169 *et seq.*
Industrial Revolution, 26, 47, 89, 304
Industry : definition, 61 ; as group of firms, 23 ; advertising by, 178
Inequality, 3, 105-6, 140-2, 216, 242-4, 254, 256-7, 415 ; *See also* Class Divisions
Inflation and deflation, 250, 294-6, 303*n*, 320, 340 *et seq.*, 345, 351*n*, 362-5, 389, 400
Information : cost of, 62-3 ; productivity of, 241, 275-6
Inheritance, 418
Insurance, 96, 260, 358 ; schemes, 142*n*, 410 ; and profit, 276 ; and risk-taking, 75-6
Integration : industrial, 127-35
Interest : rate of, chap. 21 *passim*, 19, 26, 355, 363-6, 379, 410-11 ; differences in rates of ; 258-60 ; forces governing, 268-70, 390 ; real and nominal, 253-4 ; private and social, chap. 18 *passim*, 16, 61, 106-9, 209
Invention, 25, 34, 43, 62, 89, 231, 241, 267, 405, 418*n*
Investment, 38, 142 ; and employment, 382-5 ; and risk, 93 *et seq.* ; home and foreign, 17, 36, 280, 288, 359, 374, 398-9 ; *See also* Capital Outlay, National Investment Board
Investment intermediaries, 58, 95-6, 295
Investment Trusts, 96, 102, 260, 295
Issuing Houses, 43

James, William, 44
Jenks, L. H., 100*n*
Jevons, W. S., 160, 166, 323, 382
Joint-Stock Companies, 16, 81, 90 *et seq.*, 101-3, 276-7
Joplin, Thomas, 306
Joslyn, C. S., 142*n*
Judgment, 22, 24, 44, 77

Keynes, Lord, 100*n*, 382, 385 *et seq.*
Kitchin, A. W. M., 72*n*
Kreisler, Fritz, 21

Labour : as a factor of production, 22, 24 ; and value, 148-9 ; market, 24, 54 ; *See also* Specialisation, Wages, Trade Unions
Labour Exchanges, 241
Lace-making, 58
Laissez-faire, chap. 29 *passim* ; *See also* Capitalism, Private Enterprise
Land : as a factor of production, 22, 26 *et seq.* ; marginal, 31, 196, 221 ; value of, 17 ; *See also* Rent
Landlords, 27, 35
Lavington, F., 78*n*
Law : economic, 10 ; Gresham's, 303*n* ; of diminishing marginal

productivity, 225 ; of diminishing returns, 32, 192 *et seq.*, 284 ; of diminishing utility, 160, 165
Laws of nature, 148, 396
Lebensraum, 286
Legal Tender, 297, 299, 330
Levy, H., 121*n*, 131*n*
Liberal Compromise, 415
Liefmann, R., 122*n*
Light Castings, 84
Limited Liability, 90
Liquidity, 293, 300-1 ; -preference, 37, 251, 258, 346, 386 ; and Stock Exchange, 94
Localisation of Industry, chap. 5 *passim*, 67, 83-5, 280
London Clearing House, 310, 331
Lloyd George, D., 106
Lucas, A. F., 124*n*
Luxuries, 158

Macgregor, D. H., 13
Machinery, 25, 28, 30, 36 ; and unemployment, 51, 381 ; and working conditions, 45 *et seq.* ; as stored up labour, 262 ; renewal of, 199-200
Management, 25, 68, 78, 142 ; and ownership, 101 *et seq.*; and size of firm, 80 *et seq.*; earnings of, 271 ; monetary, 365-70 ; professional, 105 *et seq.* ; *See also* Planning, Directors
Malthus, Rev. T. R., 217
Margin, 160 *et seq.* ; of cultivation, 31 ; of transference, 31, 186, 196
Marginal : cost, 172-7, 184-6, 353 ; disutility of labour, 218 ; earnings, 417 ; firm, 185, 201 ; land, 31, 196, 221 ; net product, 228 ; preference, 167, 219 ; prime cost, 412 ; product, 222 *et seq.* ; productivity, 222 *et seq.*, 233, 265-7, 390, 409 ; purchasers, 163 ; revenue, 173, 175, 198, 208, 245 ; uses, 162 ; utility, 161
Market : definition, 23 ; density, 56, 84 ; imperfection, 16, 60, 61, 73, 80, 83, 109, 154-5, 172, 180-1, 192-4, 211-2, 260, 397 ; and location of industry, 55 *et seq.*; and use of money, 294-5 ; for oranges, 152-4 ; as co-ordinator, 397 ; sharing the market, 120-1 ; spoiling the market, 179, 184 ; *See also* Open Market Operations
Marketing, 72-4, 79, 129
Markets : commodity, 276
Marquand, H. A., 117*n*, 133*n*
Marshall, Alfred, 51*n*, 83, 133*n*, 150, 284
Marx, Karl, 149, 262
Means, G. C. ; *See* Berle, A. A.
Meat Packing, 67, 84, 132 ; *See also* Slaughterhouses
Menger, Karl, 160
Mercantilism, 398-402
Migration : of industry, 55 ; of labour, 59
Milk Marketing Boards, 119, 122, 186, 207
Mill, John Stuart, 27
Miller, Margaret, 100*n*
Milling Industry, 23
Minimum standard, 418
Mining, 27 ; *See also* Coal-mining and Gold-mining
Mobility : of capital, 261, 374 ; of labour, 238-42, 372, 377-8, 380-1, 411 ; of resources, 159, 191, 280 ; social, 243-4, 393
Money : and capital, 36 ; and debts, 294 ; and income, 42 ; and liquidity, 258 ; and planning, 295 ; and welfare, 13 ; at call, 311 ; bank money, 300-2 ; paper money, 294, 297-300 ; token money, 297 ; of account, 294 ; market, 305 ; functions of, 292-5 ; implications of, 1-4 ; demand for, 343 *et seq.* ; supply of, chap. 25 *passim* ; fluctuations in value of, chap. 26 *passim*, 17, 36 ; quantity theory of, 341 *et seq.*
Monopoly, chap. 9 *passim*, 42, 62, 73, 75, 89, 143, 150, 155, 177, 201, 277-8, 353-4, 395, 401, 403, 421 ; and consumers' choice, 6 ; and distribution, 187, 232 *et seq.* ; and growth, 107 *et seq.* ; and land, 28 ; and law, 186 ; and raw materials, 186-7 ; and Trade Unions, 248 ; degrees of, 177, 211 ; discriminating, 416 ; state control of, 414 ; *See also* Profits, Restrictive Practices, Nationalisation

Monotony, 50
Motor-car Manufacturing, 3, 23, 56, 57, 59, 64, 66, 76, 131, 155, 189
Mun, Thomas, 37, 402

Napoleon, 82
National Debt, 16, 40, 212, 303-4, 321, 393, 398, 407, 410
National Health Insurance, 235, 419
National Income, 18-20, 40, 209, 226-7, 231-2, 235, 250, 344, 348, 356, 385, 393
National Investment Board, 101, 411
National Output, *See* National Income
National Wealth, 17, 36, 140
Nationalisation, 108, 136 *et seq.*, 143, 404-6, 414
Natural Rights, 396
Newcastle Vend, 121
New Deal, 414
Nickel, 60, 186
Northcliffe, Lord, 83*n*
Nuffield, Lord, 59*n*, 69

Oakeshott, W., 51
Occupations : grouping, 24 ; risky, 26
Oil, 23, 56, 132
Oligopoly, 183
Open market operations, 324, 329-30, 365-6, 390
Orange-market, 159
Ores, 130-1 ; and location, 55-6
Organisation : as a factor of production, 25 ; industrial, Part 2 *passim*
Output : ideal, 4, 143, 209 *et seq.* ; limitation of, 120 ; *See also* Monopoly
Overdraft, 307-8, 382
Overhead costs, 112, 150, 173*n*, 206 ; *See also* Cost
Overproduction, 380

Paish, F. W., 368*n*
Par value of shares, 92
Parity : mint, 36 ; purchasing power, 370-3
Parkinson, H., 97*n*
Partnership, 90, 95
Payments : Balance of, 358
Pencils : production of, 174-6
P.E.P., 134*n*
Period of Production, 93 ; long, 197 ; short, 197 *et seq.*
Philosophy : and economics, 8
Physiocrats, 396
Pigon, A. C., 404*n*
Planning, 4, 6, 25, 61, 71, 119, 137, 138-9, 295, 340, 402, 405, 408, 420-1
Population, 29, 30, 34, 43, 217, 230 ; ageing of, 239, 241, 378 ; mercantilist view of, 399 ; trend of, 267-8, 392
Post Office Savings Bank, 98*n*
Potatoes, 159, 197, 264
Pound Sterling, 294
Poverty, 2, 13, 415, 421
Power : electric, 16, 47, 77, 84 ; hydro-electric, 64, 230, 265 ; water, 47 ; and location, 54 *et seq.*
Price Discrimination, 146, 207-8, 211-2
Price Fixing, 120-1, 145
Prices : equilibrium, 152 ; fair, chap. 18 *passim*, 8, 145-6 ; fluctuations in, chap. 26 *passim*, 17, 37 ; function of, chap. 11 *passim*, 139 ; minimum, 119 ; wholesale, 339-40 ; *See also* Demand, Supply, Margin
Prime cost, 199 ; loss, 201 ; profit, 201
Printing and Publishing, 65, 67, 133
Priority, 138-9
Producers' Surplus, 278
Production : definition, 21 ; factors of, chap. 3 *passim*, 22 ; period of, 263 ; roundabout, 39 ; and prices, 356 ; *See also* Capital, Cost of Production, Supply
Productive and Unproductive Labour, 21, 222
Profit-motive, 11, 420
Profits, 28, 34 ; function of, 6, 16, 139, 210, 420 ; origin of, 272 ; gross and net, 271 ; monopoly, 187, 277 ; undistributed, 96, 117-8, 234, 256, 308 ; windfall, 29, 296, 356 ; and economies, 62 ; and enterprise, 26 ; and expansion, 88 ; and judgment, 184 ; and progress, 275 ; and taxation, 417
Promoters : Company, 43
Property : *See* Capital and Wealth

Prospecting, 26-7
Protection, 289, 399, 413
Psychology, 1, 5, 10
Public Corporations, 403-4
Public Utilities, 36
Public Works : and unemployment, 406-10
Purchasing Power : deficiency of, 12, 41-2
Pyramiding : of Joint-Stock Companies, 101-3

Quantity Theory of Money, 36-7, 400
Quasi-rent, 28 243

Racketeering, 2, 15, 110-1
Railways, 15, 16, 23, 36, 43, 62, 77, 80, 89, 97, 103, 109, 125, 133, 187, 230, 393-4
Rates, 19
Rationing, 6 ; of resources, 143
Rearmament, 203-4
Religion, 4
Rent : and cost, 29 ; and foreign trade, 291 ; and prices, 356 ; economic, 227 ; institutional, 243-4 ; urban, 221 ; of ability, 242 ; of houses, 12, 392 ; of land, 26 *et seq.* ; *See also* Cost of Production, Land
Research : and scale of production, 70 ; institutes, 63
Resourcefulness of supply, 284
Restraint of Trade, 120, 124
Restrictive practices, 112-3, 248, 414
Retailing, 72-4, 80, 86*n*, 111, 129, 185*n*, 270
Returns : diminishing and increasing, 32, 192 *et seq.*
Ricardo, David, 221
Risk : and risk-taking, 26, 48-9, 60-1, 75 *et seq.*, 93 *et seq.*, 99, 142, 270 *et seq.*, 295, 383-4, 397, 405, 417
Roads, 3, 15, 36, 62
Robbins, Professor L., 7
Robertson, Professor D. H., 39
Robinson Crusoe, 8, 37, 43, 263
Robinson, E. A. G., 59*n*, 66*n*
Roosevelt, President, 414
Ross, E. A., 75*n*
Rothamsted, 33
Rubber, 135, 159
Russia : distribution of income in 415

Salvage, 67
Satiability of Wants, 4
Savings, 5, 29, 42, 89, 94 *et seq.*, 96, 410 ; and banking, 325-6 ; and employment, 269, 382, 385-9 ; and taxation, 256, 417-8 ; function of, 38-9 ; *See also* Capital, Interest, Profits, undistributed
Scarcity : and economy, 3 ; and rent, 27 *et seq.* ; and substitution, 191 ; and value, 148-9, 217
Screws, 46, 66
Seasonal fluctuations, 376 ; loans, 313-4, 321
Sectionalism : danger of, 48
Seigniorage, 361
Self-sufficiency, 2, 43, 286
Services : and income, 20 ; and location, 60 ; *See also* Social Services
Shipbuilding, 23, 53, 57, 65, 189, 237
Shipping, 120 ; earnings, 358
Site-value, 27-8, 54, 59
Slaughterhouses, 63, 84
Slichter, Professor S. H., 45*n*, 63*n*, 70*n*
Smallholdings, 72
Smith, Adam, 27, 44, 46, 119, 400, 401, 416
Smith, Henry, 129
Smoking : economic aspects, 2 ; moral aspects, 8 ; political aspects, 10 ; psychological aspects, 5
Soap-making, 122, 124
Social Services, 3, 19, 146, 211-2, 217, 418-20
Specialisation, chap. 4 *passim*, 1, 8 ; and international trade, 279, 282 ;

and money, 292, 295 ; and scale of production, 61 ; functional, 68-72 ; in risk-bearing, 94 *et seq.*, 276
Speculation, 11, 99 *et seq.*, 344-6
Stamp, Lord, 17, 36, 255
Standardisation, 86-8, 123
State Intervention, 2, 105, 143, 210-1, chap. 29 *passim* ; *See also* Nationalisation, National Investment Board
State Investment, 265-6 ; *See also* National Investment Board
State Planning Commission, 4, 6, 138
Statistics : importance of, 12
Steel Industry, 23, 28, 45, 53, 55, 63, 67, 73, 78-9, 80, 128-9, 131, 133, 189, 192, 203
Stock Exchange, 58, 88, 99-101, 256, 295, 410
Stock-in-trade, *See* Capital, Trade
Stocks : and fluctuations, 130, 197
Stocks and Shares, 16, 39, 74-5, 91 *et seq.*, 315, 345, 411
Strawberries : market for, 197-8
Strikes, 73*n*, 77, 247
Subsidies, 19, 210, 413, 418
Substitution : at the margin, 164, 224 *et seq.* ; of factors of production, 192 ; *See also* Electricity, Scarcity
Supply : and demand, Part 3 *passim*, 3, 143, 151, 216 ; conditions, 151 ; curves, 151 ; price, 219 ; joint and composite, 202 ; *See also* Demand, Equilibrium, Price

Taussig, F. W., 142*n*
Tariffs, 8, 37, 115, 280, 371, 377 ; *See also* Protection, Trade, international
Tawney, R. H., 105*n*
Taxation, 8, 19, 140, 142,3 93, 396, 413-4, 415 ; ability theory of, 416 ; benefit theory of, 416*n* ; progressive, 415 *et seq.*
Terms of Trade, 283 *et seq.*
Textile Industries, 53, 63 ; *See also* Artificial Silk, Cotton, Wool
Thrift, 386, 412 ; and sacrifice, 219 ; incentives to, 255 ; *See also* Capital, Interest, Savings
Time-preference, 258 ; *See also* Thrift
Tinplate manufacturing, 56, 64-5, 108, 133
Trade : balance of, 37, 357
Trade Cycle, 284, 378 ; *See also* Booms and Slumps, Unemployment
Trade, International : chap. 22 *passim*, 44*n*, 52 ; finance of, 311-2 ; free, 402 ; multilateral, 285 ; effect on distribution of income, 288 ; and gold standard, 373-4 ; and public works, 409
Trade Returns, Board of, 358
Trade : terms of, 283 *et seq.*
Trade Unions, 247 *et seq.*, 364
Trading Estates, 61
Training Centres, 241
Transfer Cost, 190
Transfer Payments, 19
Transport : technical coefficients in, 191
Transport Costs, 83-5, 115
Transport Revolution, 284
Treasury Bills, 259, 365, 368
Treasury and Monetary Policy, 319, 323
Trees of the Forest, 83, 136*n* ; *See also* Whitehead
Truptil, R. J., 312*n*
Trusts, 122 *et seq.*

Uncertainty ; *See* Risk
Underemployment, 376
Understandings : and industrial combination, 119-21
Underwriting, 318, 358
Unearned Increment, 29-30
Unemployment, 2, 4, 8, 13, 24, 26, 37, 39, 41, 42, 49, 137, 146, 212, 216, 223, 347, 398-400 ; cyclical, 378 ; long-term, 60-1, 377 ; nor-

mal, 376-7; structural, 377-8; technological, 51-2; voluntary, 375, and cheap imports, 289; and distribution of income, 235; and gold standard, 364; and mobility, 241-2; and wages, 248; causes of, 382 *et seq.*; mercantilist's views of, 402; remedies for, 406 *et seq.*
Unemployment Assistance Board, 234
Unemployment Insurance, 419
Unproductive Labour, 21
Utility: defined, 13; diminishing, 160; and production, 21; *See also* Demand, Margin, Value, paradox of

Value: defined, 147; cost of production theory of, 149-50; labour theory of, 148-9; paradox of, 148, 161; *See also* Cost, Price
Velocity of Circulation, 341 *et seq.*

Wages, 2, 21, 26, chap. 20 *passim*; fair, 240-1; living, 11; nominal, 236; real, 41, 235-6; women's, 240; economy of high, 246; long run influences on, 230 *et seq.*; plasticity of, 364-5; subsistence theory of, 217; and foreign trade, 283 *et seq.*, 289, 372, 399; and invention, 231-2; and prices, 353, 355-6; and real costs, 218; cuts and employment, 41, 223*n*, 247, 412-3; disputes, 250; *See also* Strikes, Trade Unions, Unemployment
Waiting, 263
Wallpaper manufacturing, 23, 128*n*
Walras, Leon, 160
Wants: coincidence of, 292; *See also* Demand
War, 16, 19, 214, 398-9, 401; as cure for unemployment, 409; experience, 405; loans, 265; ships, 21, 93
Water: as an economic good, 14, 161-2
Watt, James, 46
Wealth: definition, 13; and capital, 36-7; and interest, 254; and welfare, 15; *See also* Capital, National Wealth
Wedgwood, J. C., 254*n*
Weights in index numbers, 334-6
Welfare, 13, 15, 212-5
Whitehead, A. N., 136*n*
Wireless Manufacturing, 59, 66, 190
Women Workers, 59, 146; wages of, 240
Wool, 57, 81

Young, Allyn, 133*n*
Young, G. M., 401*n*